Een keuze-framework voor outlier-detectiemethoden voor declaratiefraudescenario's in de Zorg

Afstudeertraject BPMIT (IM9806)

"Fraud control is a miserable business.

Failure to detect fraud is bad news; finding fraud is bad news too"

Prof. Dr. Malcolm K. Sparrow

Cursus	: IM9806 Afstudeertraject Business Process Management and IT
Student	: Maikel Maasakkers
Identiteitsnummer	: 850916439
Datum rapport	: 7 februari 2017
Datum presentatie	: 10 februari 2017

Een keuze-framework voor outlier-detectiemethoden voor declaratiefraudescenario's in de Zorg

Afstudeertraject BPMIT (IM9806)

A choice framework for outlier detection methods for healthcare claim fraud scenarios

Graduation BPMIT (IM9806)

Opleiding	: Open Universiteit, faculteit Management, Science & Technology Masteropleiding Business Process Management & IT
Programme	: Open University of the Netherlands, faculty of Management, Science & Technology Master Business Process Management & IT
Cursus	: IM9806 Afstudeertraject Business Process Management and IT
Student	: Maikel Maasakkers
Identiteitsnummer	: 850916439
Datum	: 7 februari 2017
Afstudeerbegeleider	: Dr. Ir. Ella Roubtsova
Meelezer	: Prof. Dr. Ir. Rob Kusters
Versie nummer	: 1.2
Status	: final

Voorwoord

Het rapport dat voor u ligt is het resultaat van mijn afstudeeronderzoek dat is uitgevoerd voor het afronden van de opleiding Business Process Management & IT van de Open Universiteit.

In mijn reguliere werkomgeving, kom ik veel bij zorgverzekeraars. Declaratiefraude is een groot probleem voor deze zorgverzekeraars, maar vooral ook van de maatschappij omdat de zorgkosten ieder jaar stijgen en zorgfraude hierbij naar verwachting een grote rol speelt. Met dit onderzoek probeer ik, met mijn passie voor data science, een steentje bij te dragen om zorgfraude beter te bestrijden.

De afgelopen jaren, zijn er verschillende mensen die mij gesteund hebben om het traject tot een goed einde te brengen. In het bijzonder wil ik mijn afstudeerbegeleiders Dr. Ir. Ella Roubtsova en Prof. Dr. Ir. Rob Kusters bedanken, voor de begeleiding, de waardevolle feedback, maar vooral ook het geduld. Ook wil ik Drs. Ineke Heil bedanken voor de ondersteuning en het duwtje in de rug, op de juiste momenten.

Tenslotte wil ik mijn vriendin Mariëlle en mijn kinderen Evi en Jorn enorm bedanken voor steun, tijd en ruimte die ik heb gekregen om de opleiding af te ronden.

Maikel Maasakkers

Januari 2017

Samenvatting

De kosten van de zorg stijgen ieder jaar. Volgens het CBS, waren de totale zorgkosten in 2015 € 95,3 miljard. Om de kosten beheersbaar te houden, is het van belang om te kijken hoe deze kosten verminderd kunnen worden, zonder dat dit ten koste gaat van de kwaliteit.

Volgens het rapport The Financial Cost of Healthcare Fraud en het Openbaar ministerie, ligt het fraude percentage in de zorg op minimaal 3%, waarschijnlijk meer dan 5% en mogelijk meer dan 10%. Dat zou voor de zorg neerkomen op minimaal 2 miljard euro per jaar. Zorgverzekeraars constateerden in 2015 slechts 11,1 miljoen euro aan fraude, wat dus naar alle waarschijnlijkheid het topje van de ijsberg is.

Zorgverzekeraars gebruiken momenteel voornamelijk materiële controles op declaratieniveau om declaratiefouten en –fraude op te sporen, maar maken nog nauwelijks gebruik van moderne datamining methoden om proactief frauduleuze zorgaanbieders op te sporen.

Het doel van het onderzoek is een keuze-framework, voor de selectie van outlier-detectiemethoden voor declaratiefraudescenario's in de Zorg. Het onderzoek volgt de Design Science methode , waarbij de volgende stappen doorlopen worden: identificatie en motivatie van het probleem, eisen aan de oplossing, ontwerp en ontwikkeling, demonstratie, evaluatie en communicatie.

Op basis van kenmerken van 62 bekende fraudescenario's, zijn zes generalisaties opgesteld. Per generalisatie, zijn op basis van de literatuur, de mogelijk toepasbare outlier-detectiemethoden bepaald, waaronder: regressie, boxplot en denity-based clustering. De methoden zijn uitsluitend unsupervised, omdat er onvoldoende bekende fraudegevallen bekend zijn.

Voor de verschillende scenario generalisaties zijn test datasets gegenereerd waarmee de toepasbaarheid van de methodes per generalisatie zijn geëvalueerd. De methodes zijn vergeleken op basis van de *recall*, *precision* en f1-score, welke per methode afleidbaar zijn uit de convolutie matrix.

Het onderzoek heeft een concreet keuze-framework opgeleverd, waarmee op basis van een beperkt aantal kenmerken van fraudescenario's, de meest toepasbare methoden afgeleid kunnen worden. Outlier-detectiemethoden, hebben vaak meerdere varianten met specifieke voor- en nadelen. De belangrijkste verschillen tussen deze varianten, worden in het framework ook toegelicht.

Summary

Health care costs rise every year. According to the Dutch central bureau of statistics (CBS), total healthcare costs in 2015 were 95,3 billion Euro. To keep these costs under control, it is important to explore options for cost reduction, however not at the expense of quality.

According to the report 'The Financial Cost of Healthcare Fraud' and the Public Prosecution Service (OM), the minimal level of healthcare fraud is 3%, likely more than 5% and possibly more than 10%. This would mean a minimum cost of 2 billion Euro per year. Healthcare insurers detected in 2015 only 11.1 million Euro fraud costs, suggesting this may be the tip of the iceberg.

Healthcare insurers currently mainly use material checks of expense statements to detect errors as well as fraud, whereas they hardly use modern datamining methods to pro-actively detect fraudulent healthcare providers.

The purpose of this research is to provide a framework to select outlier detection methods to detect expense statement fraud scenario's in Healthcare. This research adopts the Design Science process, which includes six steps: Problem identification and motivation, definition of the objectives for a solution, design and development, demonstration, evaluation, and communication.

On the basis of 62 characteristics of well-known fraud scenarios, six generalizations are determined. Per generalizations, possible outlier detection methods are determined by literature review, including: Regression, boxplot and density-based clustering. These are purely unsupervised methods, because there are not enough known fraud cases.

For each of the scenario generalizations, test datasets are generated to evaluate the applicability of the outlier detection methods per scenario. Methods are compared with respect to recall, precision and F1-score, which can be derived from the convolution matrix.

This research has resulted in a concrete framework to derive the most applicable methods to detect outliers on the basis of a limited number of fraud scenario features. Outlier detection methods often have multiple variants with specific characteristics, pros and cons. The most important differences between these variants are also highlighted within the framework.

Leeswijzer

Acroniemen en afkortingen

Acroniem / afkorting	Omschrijving
ANW-toeslag	Avond, Nacht en Weekend toeslag
AWBZ	Algemene Wet Bijzondere Ziektekosten
DBC	Diagnose Behandel Combinatie
CIZ	Centrum Indicatiestelling Zorg
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
DeBaCl	DEnsity-BAsed CLustering
FIOD	Fiscale Inlichtingen en Opsporingsdienst
FTE	Fulltime Equivalent
GMM	Gaussian Mixture Models
IGZ	Inspectie voor de Gezondheidszorg
IKA	InterKwartielAfstand
kNN	k-Nearest Neighbors
NZa	Nederlandse Zorgautoriteit
OM	Openbaar Ministerie
RANSAC	RAndom SAmple Consensus
SHA	System of Health Accounts
SZW	Sociale Zaken en Werkgelegenheid
VWS	ministerie van Volksgezondheid, Welzijn en Sport
OM	Openbaar Ministerie
SHA	System of Health Accounts
SZW	Sociale Zaken en Werkgelegenheid
ZN	Zorgverzekeraars Nederland

Tabel 1.1: Acroniemen en afkortingen

Begrippen

Begrip	Omschrijving
Fraude	Het opzettelijk overtreden van een wet, regel of voorwaarde, waardoor een onterecht voordeel (zoals vergoeding of dekking) wordt behaald. *Bron: Kenniscentrum Fraudebeheersing*
Zorgaanbieder	De natuurlijke persoon of rechtspersoon die beroeps- of bedrijfsmatig zorg verleent, als bedoeld in artikel 1, onderdeel c, onder 1°, Wmg. Bijvoorbeeld Tandarts, Apotheek en Ziekenhuis. *Bron: NZa, BELEIDSREGEL CA-BR-1604*
Zorgfraude	Van fraude in de zorg wordt gesproken, indien sprake is van opzettelijk gepleegde onrechtmatige feiten, die ten laste komen van voor de zorg bestemde middelen. Bij fraude moet voldaan zijn aan de volgende elementen: ❑ (Financieel) verkregen voordeel; ❑ Overtreden van declaratieregels; ❑ Opzettelijk en misleidend handelen. *Bron: Regiegroep 'verbetering van zorgfraudebestrijding'*

Tabel 1.2: Begrippen

Gebruikte symbolen

Symbool	Omschrijving
#	Aantal
Ø	Gemiddelde
%	Percentage

Tabel 1.3: Gebruikte symbolen

Inhoudsopgave

1. Introductie

1.1. Inleiding

De kosten van de zorg stijgen ieder jaar. Om de kosten beheersbaar te houden, is het van belang om te kijken hoe deze kosten verminderd kunnen worden, zonder dat dit ten koste gaat van de kwaliteit.

Zorgverzekeraars gebruiken momenteel voornamelijk materiële controles op declaratieniveau om declaratiefouten en –fraude op te sporen. Verschillende zorgverzekeraars hebben hier ook fraudeteams op zitten, maar zij analyseren voornamelijk reactief. Bijvoorbeeld op signalen van de afdelingen declaratieverwerking of op basis van signalen van patiënten over specifieke zorgaanbieders. Het keuze-framework biedt data-analisten richtlijnen om nieuwe datamining technieken proactief toe te passen bij het opsporen van frauduleuze zorgaanbieders.

Volgens het rapport The Financial Cost of Healthcare Fraud (Gee & Button, 2015) en het Openbaar ministerie, ligt het fraude percentage in de zorg op minimaal 3%, waarschijnlijk meer dan 5% en mogelijk meer dan 10%. Dat zou voor de zorg neerkomen op minimaal 2 miljard per jaar. Zorgverzekeraars constateerden in volgende (Zorgverzekeraars Nederland, 2017) in 2015 slechts 11,1 miljoen euro aan fraude, wat dus naar alle waarschijnlijkheid het topje van de ijsberg is.

Het doel van het onderzoek is een keuze-framework, voor de selectie van outlier-detectiemethoden voor declaratiefraudescenario's in de Zorg. Het onderzoek volgt de Design Science methode, waarbij de volgende stappen doorlopen worden: identificatie en motivatie van het probleem, eisen aan de oplossing, ontwerp en ontwikkeling, demonstratie, evaluatie en communicatie.

Vanwege de privacygevoelige aard van de declaratiedata, was het niet mogelijk om productiedata te gebruiken voor het onderzoek. Daarom is er in het onderzoek veelvuldig gebruik gemaakt van domein experts van Truston B.V. en twee zorgverzekeraars.

In hoofdstuk 1, wordt een overzicht gegeven van de context en de gebruikte begrippen in het onderzoek. Tevens wordt de wetenschappelijke - en maatschappelijke relevantie toegelicht en zowel de probleemstelling als de opdracht geformuleerd. Hoofdstuk 2 beschrijft de aanpak en de uitvoering, de resultaten en conclusies van de literatuurstudie.

Hoofdstuk 3 geeft het doel van het empirisch onderzoek en in hoofdstuk 4 wordt de toegepaste methode toegelicht aan de hand van het onderzoeksmodel, de onderzoeksstrategie en de onderzoeksaanpak.

De uitvoering van het onderzoek, wordt in hoofdstuk 5 beschreven. Hier wordt de opzet van het keuze-framework, de kenmerken en classificatie van de fraude scenario's besproken. Per generalisatie van fraudsescenario's, worden de resultaten van de verschillende methoden weergegeven en toegelicht aan de hand van metrieken.

De resultaten van het onderzoek, als mede het uiteindelijke keuze-framework, worden in hoofdstuk 6 besproken. Hoofdstuk 7 belicht de kanttekeningen en de concessies, die in het onderzoek gedaan zijn.

In hoofdstuk 8 staan de conclusies en aanbevelingen. Tenslotte volgt in hoofdstuk 9, de reflectie over de kwaliteit van het onderzoek en het proces.

1.2. Context

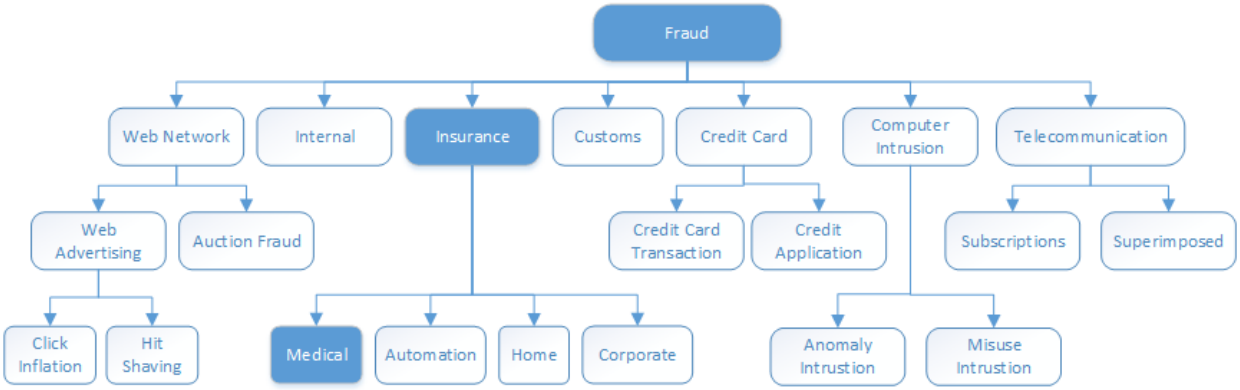
Dit onderzoek gebruikt terminologie van zowel het domein Zorg als Data Science. De belangrijkste begrippen worden ik dit hoofdstuk toegelicht.

1.2.1. Fraude

Het kenniscentrum Fraudebeheersing definieert fraude als volgt:

> *"Het opzettelijk overtreden van een wet, regel of voorwaarde, waardoor een onterecht voordeel (zoals vergoeding of dekking) wordt behaald."*

Fraude komt voor in diverse domeinen en in verschillende verschijningsvormen. Een taxonomie voor fraude typeringen, wordt gegeven door (Laleh & Azgomi, 2009):



Tabel 1.1: Taxonomie fraude types

Dit onderzoek heeft betrekking fraude binnen *Medical Insurance* ofwel *Health Insurance* (Zorgverzekeringen).

1.2.2. Zorgfraude

De regiegroep 'verbetering van zorgfraudebestrijding' hanteert voor de definitie van zorgfraude:

> Van fraude in de zorg wordt gesproken, indien sprake is van opzettelijk gepleegde onrechtmatige feiten, die ten laste komen van voor de zorg bestemde middelen. Bij fraude moet voldaan zijn aan de volgende elementen:
> - ❑ (Financieel) verkregen voordeel;
> - ❑ Overtreden van declaratieregels;
> - ❑ Opzettelijk en misleidend handelen.

Bij zorgfraude zijn bewust wettelijke regels overtreden, anderen zijn misleid en de dader heeft er financieel voordeel bij. Een patiënt heeft bijvoorbeeld een rekening gekregen van een zorgaanbieder waar deze nooit geweest is of er is meer zorg in rekening gebracht dan geleverd.
Indien er sprake is van niet opzettelijke onterechte declaraties, spreekt men niet van fraude. Het is vaak lastig aan te tonen of een onterechte declaratie een fout is of daadwerkelijk moedwillige fraude.

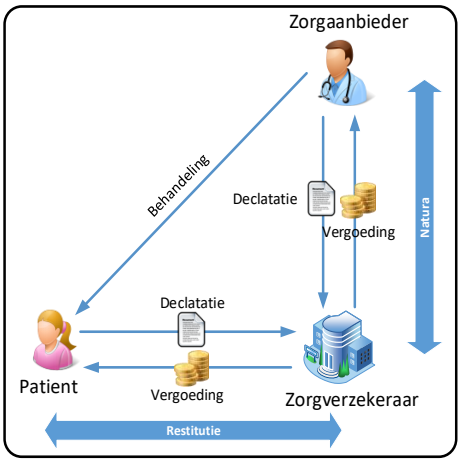
1.2.3. Declaratieproces

Restitutie / Natura

Binnen het declaratieverkeer wordt onderscheid gemaakt tussen restitutie - en natura declaraties. Bij naturadeclaraties verstuurt de zorgaanbieder de declaratie direct naar de zorgverzekeraar en betaalt de zorgverzekeraar uit naar de zorgaanbieder. De patiënt is hierbij dus niet direct bij het declaratie- en betalingsverkeer betrokken.

Bij restitutiedeclaraties declareert de patiënt (een deel van) de kosten aan de zorgverzekeraar. Indien de patiënt de declaratie van de zorgaanbieder zelf voorgeschoten heeft, dan betaalt de zorgverzekeraar het bedrag terug aan de patiënt. Anders wordt het bedrag direct aan de zorgaanbieder betaald.

Bij veel declaraties ziet de patiënt niet wat er aan de zorgaanbieder betaald wordt en kan dit ook niet controleren.



Figuur 1.1: Declaratieproces

1.2.4. Declaratiefraude scenario's

Er zijn verschillende scenario's om fraude binnen de zorg te plegen. Omdat de betalingen naar zorgaanbieders via declaraties verlopen, spreekt heeft fraude door zorgaanbieders ook wel van declaratiefraude. Het Rapport Onderzoek Zorgfraude (NZa, 2014) van de Nederlandse Zorgautoriteit, heeft de meest voorkomende fraudescenario's in de Zorg onderzocht. Bij het inventariseren van de frauderisico's op de diverse zorgmarkten is door het NZa de categorisering van (Sparrow, 2000) gebruikt. Deze categorieën zijn ook gebruikt bij de fraude fraudescenario's in bijlage 1.

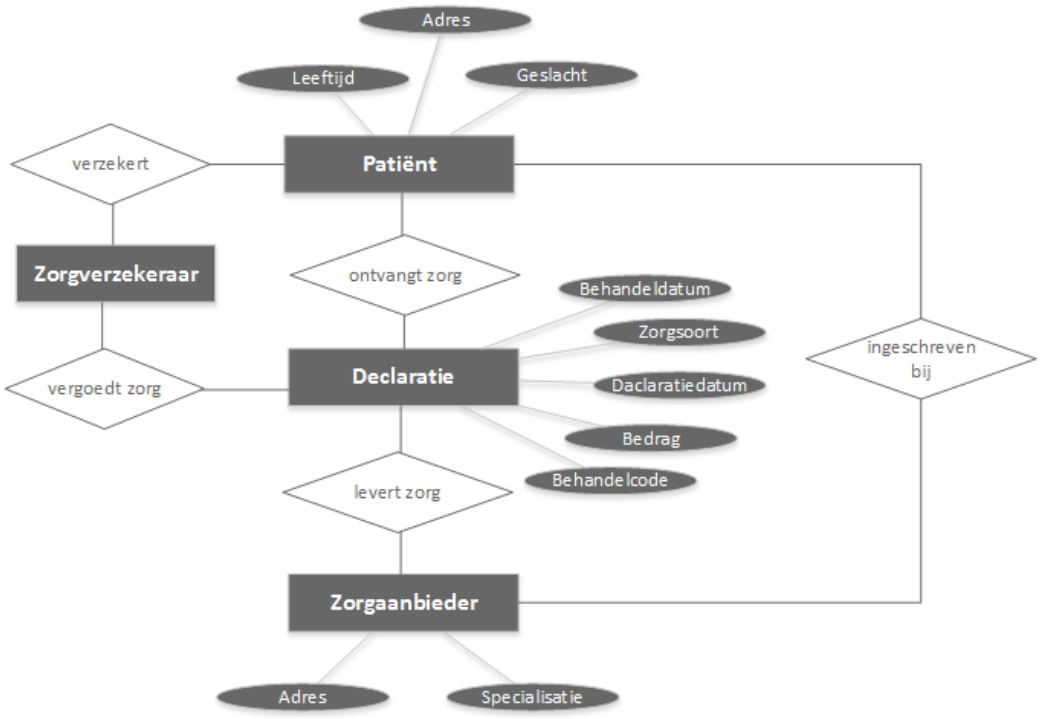
Categorie	Toelichting
Spookzorg	Bij spookzorg is zorg die gedeclareerd is niet geleverd. Dit kan door fictieve patiënten op te voeren, of behandelingen te declareren die niet hebben plaatsgevonden bij bestaande patiënten. Het laatste scenario is het meest waarschijnlijk.
Dubbel claimen	Aanbieders kunnen de zorg die zij leveren *dubbel claimen:* bijvoorbeeld bij de patiënt en verzekeraar, twee keer bij dezelfde verzekeraar of bij twee verschillende verzekeraars. Eigenlijk zou dit soort overtredingen gemakkelijk in basiscontroles van verzekeraars zichtbaar moeten worden.
Dubbele bekostiging	*Dubbele bekostiging* komt ook voor bij aanbieders met een vaste en een variabele component (bijvoorbeeld het consult van huisartsen en de M&I-verrichting chirurgie). Het risico is hier dat iets dat al in de behandeling zit via de toeslag nog een keer extra wordt gedeclareerd.
Upcoding	Bij *upcoding* declareert een aanbieder een duurdere behandeling dan de geleverde zorg rechtvaardigt. Verzekeraars kunnen dit vrijwel alleen met materiële controles boven water krijgen.

Opknippen	Frauderen door *op te knippen* kan voorkomen als een bundel van zorgactiviteiten declarabel is, maar ook de activiteiten los van elkaar
Ten onrechte laten bijbetalen	Aanbieders kunnen de *patiënt ten onrechte laten bijbetalen* voor zorg die in de geleverde prestatie hoort. Een voorbeeld hiervan is het bijbetalen voor toiletpapier in de AWBZ. Dit valt gewoon onder de AWBZ.
Meer zorg leveren dan noodzakelijk	Aanbieders kunnen hun patiënten *meer zorg leveren dan noodzakelijk*, bijvoorbeeld meer zittingen fysiotherapie dan nodig.
Onderbehandelen	Ook *onderbehandelen* kan lucratief zijn: bijvoorbeeld minder uren zorg leveren dan afgesproken in de AWBZ.
U-bochtconstructie	De *U-bochtconstructie* is een combinatie van dubbel declareren en opknippen: iemand wil een spiraaltje, en krijgt van de specialist te horen dat zij die zelf moet kopen bij de (ziekenhuis)apotheek. Het spiraaltje zelf zit echter al in de DBC. Zo ontvangt het ziekenhuis een dubbele vergoeding of een vergoeding voor niet gemaakte kosten.
Verwijsvergoeding	Een *verwijsvergoeding* is een manier om meer patiënten te krijgen, bijvoorbeeld: een ziekenhuis betaalt een verwijsvergoeding aan huisartsen. Patiënten worden niet direct gedupeerd, maar krijgen mogelijk niet de meest optimale zorg.
Onverzekerde zorg	Aanbieders kunnen onverzekerde zorg als verzekerde zorg declareren. Dit risico is er vooral in markten waar zorg onder bepaalde voorwaarden wel onder het basispakket valt, en onder bepaalde voorwaarden niet (mondzorg, fysiotherapie, medisch specialistische zorg). Deze vorm van fraude is in ieder geval voor de individuele patiënt lucratief, die krijgt de zorg hierdoor vergoed. Ook voor de aanbieder kan er winst zijn als er een verschil in prijs zit tussen de verzekerde en de onverzekerde zorg. Deze vorm van fraude is vrijwel alleen op te sporen met materiële controles.

Figuur 1.2: Fraudecategorieën

1.2.5. Gegevensmodel

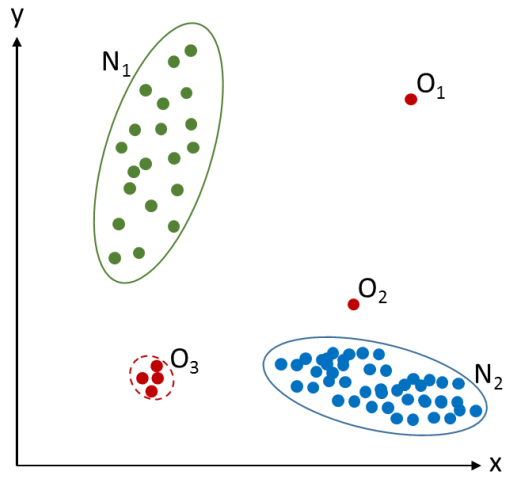
Onderstaand gegevensmodel geeft een weergave van de betrokken entiteiten en een aantal relevante kenmerken (ofwel attributen), die in het onderzoek bij de fraudescenario's gebruikt worden.



Figuur 1.3: Gegevensmodel declaratieverwerking

1.2.6. Outliers

De term *outlier* is afkomstig van de data science. Deze term is veel gebruikt in dit onderzoek. Outliers zijn patronen in data, welke niet voldoen aan de definitie van het normale gedrag, of voldoen aan de definitie van uitzonderlijk gedrag.



Figuur 1.4: Voorbeeld outliers in een 2-dimensionale dataset. N1 en N2 zijn normale regio's. O1 en O2 zijn individuele outlier instanties. O3 is een outlier regio

In de meeste dataset komen Outliers voor en zijn vaak het gevolg van (Chandola, Banerjee, & Kumar, Outlier Detection: A Survey, 2007):

- *Frauduleuze activiteiten*, zoals creditcard fraude en declaratiefraude;
- *Machinale fouten*, door bijvoorbeeld afwijkingen en slijtage in de apparatuur;
- *Verandering in de omgeving*, zoals koopgedrag of klimaat;
- *Menselijke fouten*, zoals een rapportagefout.

Typering

Type I Outliers

Een individueel afwijkende observatie wordt een type I outlier genoemd. Dit is de eenvoudigste type, waarop de meeste outlier detectietechnieken gebaseerd zijn. De observatie wijkt af van het normaal omdat de attributen inconsistent zijn met de attributen van de overige observaties.

Type II Outliers

Bij dit type outlier wijkt de observatie af omdat de observatie niet consistent is binnen een bepaalde context. Dezelfde observatie kan in een andere context wel tot de norm behoren. De context is volgordelijk of ruimtelijk van aard, bijvoorbeeld de tijd.

Type III Outliers

Dit type outlier wijst op een afwijking t.o.v. een subset van de overige observaties. Type III outliers zijn altijd data reeksen, waarbij de observaties op zich geen outliers zijn. Ook hier is de data volgordelijk of ruimtelijk van aard, bijvoorbeeld de tijd.

1.3. Relevantie

1.3.1. Wetenschappelijke relevantie

Dit onderzoek draagt bij aan de toepassing van datamining methoden in nieuwe domeinen.

Datamining methoden worden steeds breder ingezet. Naast de reguliere statistische methoden, worden de laatste jaren steeds meer machine learning methoden toegepast en nieuwe technieken ontwikkeld.

Er is veel onderzoek gedaan naar de algemene toepasbaarheid van outlier-detectietechnieken, zoals (Chandola, Banerjee, & Kumar, Anomaly Detection : A Survey, 2009) en (Agrawal & Agrawal, 2015). Er is echter nog geen onderzoek gedaan naar de toepasbaarheid binnen het zorgdomein. Het onderzoek naar zorgfraude (NZa, 2014), verwijst bij het advies om zorgfraude op te sporen vaak naar dataminig, maar geeft daarbij niet aan om welke specifieke methodes het dan gaat.

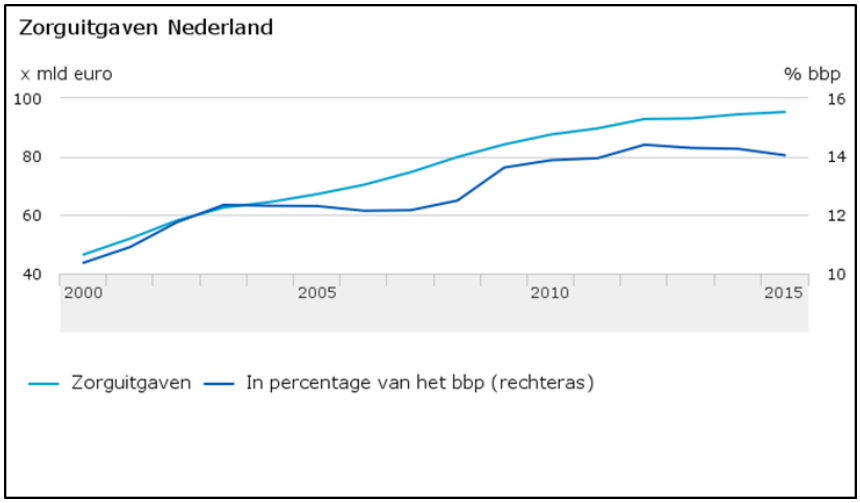
1.3.2. Maatschappelijke relevantie

Dit onderzoek draagt bij aan de toepassing van nieuwe datamining methoden voor het opsporen van declaratiefraude, waardoor de zorgkosten verlaagd worden. De kosten van de zorg stijgen ieder jaar. Om de kosten beheersbaar te houden, is het van belang om te kijken hoe deze kosten verminderd kunnen worden, zonder dat dit ten koste gaat van de kwaliteit.

Volgens het rapport The Financial Cost of Healthcare Fraud (Gee & Button, 2015) en het Openbaar ministerie, ligt het fraude percentage in de zorg op minimaal 3%, waarschijnlijk meer dan 5% en mogelijk meer dan 10%. Dat zou voor de zorg neerkomen op minimaal 2 miljard per jaar. Zorgverzekeraars constateerden volgens (Zorgverzekeraars Nederland, 2017) in 2015 slechts 11,1 miljoen euro aan fraude, wat dus naar alle waarschijnlijkheid het topje van de ijsberg is.

Kosten gezondheidszorg : € 95,3 miljard (2015)
Geschatte fraude : min. 3% (€ 2,9 miljard)
Ontdekte fraude : 0,4% (€ 11,1 miljoen)

Onderstaande grafiek laat de stijging van de zorguitgaven zien en de verhouding tot het bruto binnenlands product (bbp):



Figuur 1.5:Zorguitgaven, bron CBS, 2016

Uit het onderzoek naar zorgfraude (NZa, 2014), dat is uitgevoerd in opdracht van De Nederlandse Zorgautoriteit, worden verschillende fraudescenario's besproken, met daarbij een aantal voorbeelden. Ook worden aanbevelingen gegeven hoe men deze scenario's zou kunnen opsporen,

waarbij in veel gevallen de maatregel “datamining” wordt genoemd. Datamining is echter een erg breed begrip en biedt daarmee geen concrete handvatten om deze fraudescenario’s te onderzoeken.

1.4. Probleemstelling

De huidige moderne zorg-informatiesystemen, bieden diverse mogelijkheden om declaraties te beoordelen op legitimiteit. De controles in de systemen richten zich om een aantal specifieke probleemgebieden zoals: foutieve - en incomplete data invoer, dubbele declaraties en niet-gedekte behandelingen.

Hoewel de systemen gebruikt kunnen worden om bepaalde klassen van fraude op te sporen, zijn de fraudedetectie mogelijkheden veelal beperkt omdat de detectie gebaseerd is op eenvoudige regels die vooraf opgesteld zijn door domein experts.

Het probleem hierbij is dat er met beperkte controles, naar verwachting slechts 0,4% van de fraude ontdekt wordt.

Om declaratiefraude beter op te kunnen sporen, kan er meer gebruik gemaakt worden van geavanceerde datamining technieken voor fraude detectie. Voor zorgverzekeraars is het van belang om inzicht te krijgen in de effectiviteit van de verschillende technieken bij verschillende fraudescenario’s.

Binnen datamining, worden outlier-detectiemethoden vaak gebruikt om fraude op te sporen. Deze methoden worden echter nog niet breed toegepast voor het opsporen van declaratiefraude in de Zorg.

Voor de toepassing van outlier-detectiemethoden in de Zorg, ontbreekt er een concrete werkwijze, waarmee voor fraudescenario’s bepaald kan worden, welke methoden het meest toepasbaar zijn.

1.5. Opdrachtformulering

Doel

Op basis van de probleemstelling, is het doel als volgt gedefinieerd:

Er dient een werkwijze beschreven te worden om declaratiefraude scenario’s te classificeren en hiervoor outlier-detectiemethoden te bepalen.

Hoofdvraag

Om een werkwijze af te leiden, zal de centrale vraag beantwoord moeten worden:

Welke outlier-detectiemethoden kunnen toegepast worden op welk klassen declaratiefraude scenario’s?

Deelvragen

De volgende deelvragen zijn opgesteld om de hoofdvraag te kunnen beantwoorden:

1. Welke outlier-detectiemethoden zijn er?
2. Wat is de algemene toepasbaarheid van de outlier-detectiemethoden?
3. Op basis van welke indicatoren, kunnen de outlier-detectiemethoden vergeleken worden?
4. Welke declaratiefraude scenario’s zijn er?
5. Wat is een keuze-framework voor outlier-detectiemethoden voor declaratiefraude scenario’s in de Zorg?
6. Hoe kunnen declaratiefraude scenario’s geclassificeerd worden?
7. Wat is het motivatie voor een keuze van outlier-detectiemethoden voor de klassen declaratiefraude scenario’s?

Vragen 1 t/m 4 worden behandeld in de literatuurstudie. Het empirisch onderzoek geeft antwoord op de vragen 5 t/m 7.

1.6. Scope

De onderstaande aspecten vallen buiten scope van het onderzoek:

Type II en III outliers

Dit onderzoek richt zich uitsluitend op type I outliers. Type II en III outliers zijn ook niet expliciet benoemd in het onderzoek naar zorgfraude (NZa, 2014).

Supervised methoden

Supervised methoden vallen buiten scope omdat deze niet toepasbaar zijn of reeds toegepast worden.

Hybride methodes

Er bestaan situaties, waarbij het toepassen van meerdere outlier-detectiemethode noodzakelijk is om het gewenste resultaat te bereiken. (Agrawal & Agrawal, 2015) beschrijven verschillende hybride methodes, welke onderverdeeld kunnen worden in:

- ❑ Cascading supervised methoden: meerdere supervised methodes na elkaar uitvoeren;
- ❑ Combineren van supervised – en unsupervised methoden.

Enquêtes/evaluatie

Fraude door zorgaanbieders is vaak lastig te constateren omdat de patiënt geen inzicht krijgt in de gedeclareerde behandelingen. Een evaluatieformulier dat door de zorgverzekeraar naar de Patiënt gestuurd kan worden, is een effectief middel dat sinds 2015 vaak toegepast wordt.

Materiële controles

Voor heldere scenario's, die door relatief eenvoudige materiële controles bij zorgverzekeraars opgespoord kan worden, zijn de outlier grenzen duidelijk. Outlier-detectie is hier niet nodig.

AWBZ

AWBZ valt buiten de scope van dit onderzoek omdat het declaratieproces en de deelnemers aan het proces afwijken van de andere zorgsoorten.

Patiënt

Fraude door de patiënt valt buiten de scope van dit onderzoek.

1.7. Globale onderzoeksaanpak

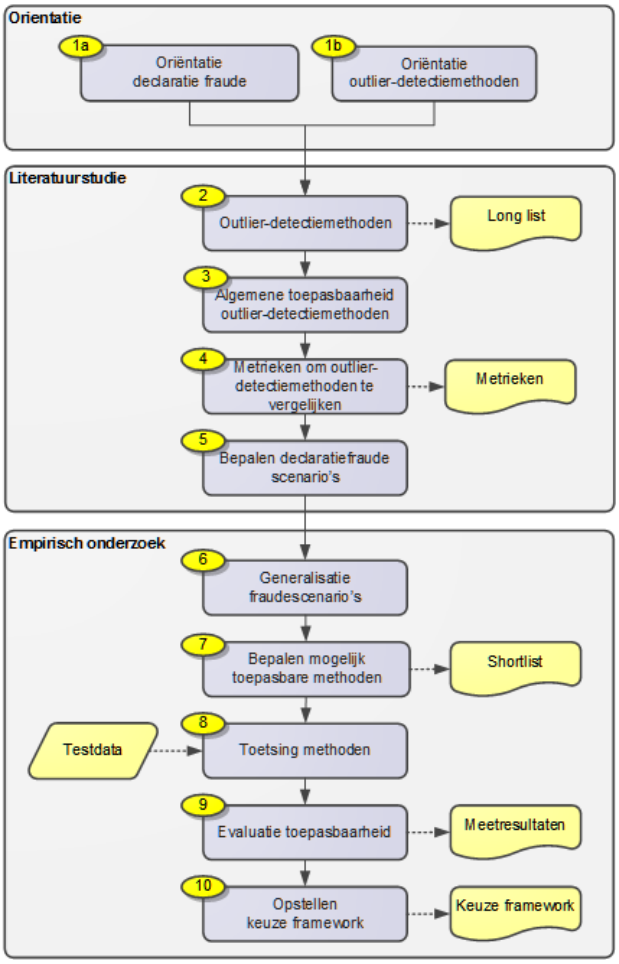
1.7.1. Design Science Methode

Het onderzoek wordt uitgevoerd volgens de Design Science methode van (Peffers, Tuunanen, Rothenberger, & Chatterjee, 2007). Deze methode sluit aan bij onderzoek waarbij een implementatie o.b.v. literatuur ontwikkeld en geëvalueerd wordt.

De fases in het onderzoek zijn:

Fase volgens Design Science	Concretisering
1. Problem identification and motivation	Onderzoeksvraag
2. Define the objectives for a solution	Eisen opstellen aan het keuze-framework
3. Design and development	Ontwerp van mogelijk toepasbare methoden
4. Demonstration	Testen met generalisaties
5. Evaluation	Evaluatie van resultaten
6. Communication	Communicatie van het keuze-framework middels een beslisboom.

Het resultaat wordt bereikt, door de onderstaande stappen te doorlopen. De beschrijving van de stappen, wordt toegelicht in onderzoeksaanpak van de Literatuurstudie en het Empirisch onderzoek.



Figuur 1.6: Onderzoeksaanpak

2. Literatuurstudie

2.1. Onderzoeksaanpak

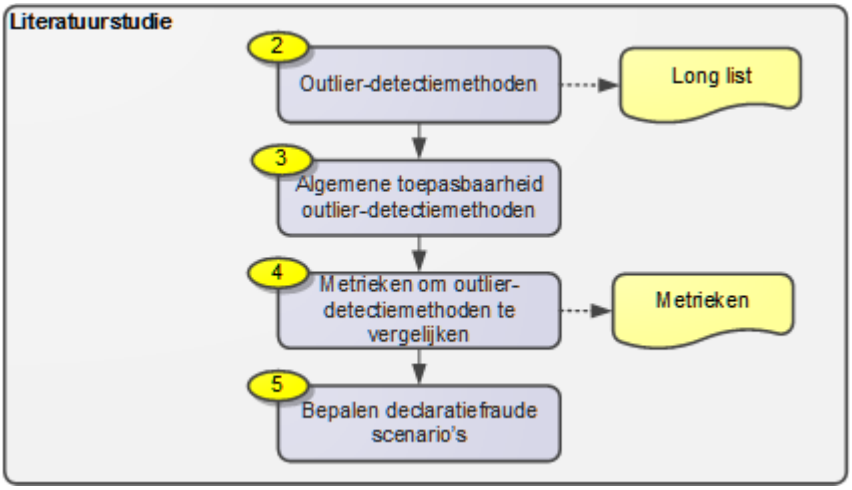
In de literatuurstudie, worden de volgende deelvragen beantwoord:

1. Welke outlier-detectie methoden zijn er?
2. Wat is de algemene toepasbaarheid van de outlier-detectiemethoden?
3. Op basis van welke indicatoren kunnen de outlier-detectiemethoden vergeleken worden?
4. Welke declaratiefraude scenario's zijn er?

Op basis van bestaande artikelen, worden outlier-detectiemethoden in kaart gebracht en de deelvragen 1 t/m 3 beantwoord.

Voor het bepalen van fraudescenario's voor vraag 4, worden domeindocumenten gebruikt.

Voor de literatuurstudie, is onderstaand stappenplan opgesteld:



Figuur 2.1: Aanpak literatuurstudie

2.2. Uitvoering

Bronnen

Voor de literatuurstudie, zijn de volgende bronnen geraadpleegd:

- ❑ Google Scholar;
- ❑ Digitale bibliotheek OU: ACM Digital Library, Springer, IEEE Digital Library, ScienceDirect (Elsevier);
- ❑ Website NZa.

Zoekwoorden

Om de relevante artikelen initieel te vinden, zijn de volgende zoekwoorden gebruikt:

Categorie	Zoekwoorden EN	Zoekwoorden NL
Fraude	▪ fraud	▪ fraud
(zorg)declaratiefraude	▪ insurance fraud ▪ healthcare fraud ▪ claims fraud	▪ declaratiefraude ▪ zorgfraude
Outlier detectie	▪ fraud detection ▪ outlier detection ▪ anomaly detection ▪ outlier detection algorithm ▪ k-nearest neighbor	

	▪ clustering	
Vergelijkingsmodellen	▪ recall ▪ percision ▪ accuracy ▪ ROC ▪ AUC	

Gevonden artikelen en relevantie

Categorie	#Gevonden	#Relevant	#Gebruikt
Fraude	100+	23	6
(zorg)declaratiefraude	50+	27	12
Outlier-detectiemethodes	100+	67	30
Vergelijkingsmodellen	100+	80+	5
Totaal			**53**

2.3. Resultaten en conclusies

2.3.1. Outlier-detectiemethoden

Outlier detectie refereert naar het probleem om patronen in data te ontdekken, welke afwijken van het normale gedrag. Outlier detectie is een onderwerp dat breed onderzocht is en toepassingen kent in verschillende domeinen, zoals: creditcard-, verzekering- en belastingfraude, intrusion detection voor cyber security, foutdetectie in bedrijfskritische systemen, detecteren van medische afwijkingen en opsporen van personen en groepen met ongewenste intenties.

De noodzaak voor Outlier detectie, ligt in het feit dat het in de meeste toepassingsgebieden betrekking heeft op waardevolle en bedrijfskritische informatie. Bij medische toepassingen kan een afwijkend patroon duiden op een ernstige ziekte en bij creditcard transacties op het gebruik van een gestolen creditcard.

Outlier detectie wordt toegepast in verschillende toepassingsgebieden, wat tot een grote diversiteit aan outlier detectietechnieken geleid heeft. Veel van deze technieken zijn ontwikkeld om een domeinspecifiek probleem op te lossen, maar er zijn ook technieken ontwikkeld die breder toepasbaar zijn.

Supervised versus unsupervised methoden

Naast de input data, is het ook mogelijk dat er gebruik gemaakt kan worden van gelabelde trainingsdata. Van de trainingsdata is dan bekend of deze onder Normaal of Outlier geclassificeerd wordt. Outlier detectie o.b.v. trainingsdata is een techniek die veel gebruik wordt bij Machine learning (Mitchell, 1997) en statistische methoden (Vapnik, 1995). Op basis van het gebruik van gelabelde traingingsdata, kunnen outlier-detectiemethoden in drie categorieën verdeeld worden:

Supervised outlier-detectiemethoden

Uitgangspunt is dat alle normale - en outlier gevallen gelabeld zijn. Deze methode probeert op basis van de trainingsset met bekende, geclassificeerde gevallen, een voorspelling te doen voor nieuwe, niet-geclassificeerde gevallen.

Het probleem bij deze techniek is dat er de data eerst gelabeld dient te worden, wat een tijd kost en vaak bij fraude maar beperkt mogelijk. Voor detectie van declaratiefraude is deze techniek dan ook niet direct bruikbaar omdat de gelabelde data niet of slechts in enkele gevallen beschikbaar is.

Semi-supervised outlier- detectiemethoden

Hierbij is een klein deel van de instanties gelabeld, maar het overgrote deel niet. Een typische benadering is om de niet gelabelde instanties te classificeren als normaal.

Unsupervised outlier- detectiemethoden

Deze methoden, gaan niet uit van gelabelde trainingsdata. Uitgangspunt hierbij is dat normale gevallen vaker voorkomen dan outliers. Instanties of clusters van instanties die niet vaak voorkomen, kunnen zo geclassificeerd worden als outliers. Clusters die vaak voorkomen, worden geclassificeerd als normaal. Deze techniek leidt vaak tot veel false positives, omdat de aannames vaak niet correct blijken.

Het onderzoek richt zich op de methoden die toepasbaar zijn voor het opsporen van declaratiefraude in de zorg. Omdat er onvoldoende fraudegevallen bekend zijn om outliers te labelen, worden uitsluitend unsupervised methoden uit de literatuurstudie besproken.

Clustering methoden

Clustering (Jain & Dubes, 1988) is een veel gebruikte leertechniek om overeenkomstige observaties te groeperen in clusters. Clustering wordt met name *unsupervised* toegepast. Op het eerste gezicht lijkt clustering niet direct toepasbaar om outliers te detecteren. Outliers kunnen bij clustering naar boven komen omdat deze niet tot een cluster behoren of een relatief klein cluster vormen.

Voorbeelden zijn: k-means, DBSCAN, DeBaCl en OPTICS

Nearest Neighbor Technieken

Nearest neighbor is een breed toegepaste methode in machine learning en datamining. Hierbij wordt een observatie vergeleken met de dichtstbijzijnde buren. Deze methode wordt o.a. toegepast bij clustering, classificatie en outlier detectie.

Om de afstand tussen observaties te bepalen wordt vaak de euclidische afstand toegepast. De Mahalanobische afstand normaliseert alle dimensies en houdt daarmee ook rekening met skewing van data. (Otey, Ghoting, & Parthasarathy, 2006) behandelt een methode om de afstand te bepalen o.b.v. categorische en doorlopende attributen.

Het bepalen of een instantie een outlier is, wordt vaak de definitie in (Knorr & Ng, 1998) gebruikt: *"A point p in a data set is an outlier with respect to the parameters k and λ, if no more than k points in the data set are at a distance λ or less from p."*

Statistische Methoden

De basis voor statistisch methodes om outliers te detecteren, ligt in de definitie van (Anscombe & Guttman, 1960): *"An outlier is an observation which is suspected of being partially or wholly irrelevant because it is not generated by the stochastic model assumed."*

Volgens deze definitie kunnen outliers gedetecteerd worden door het opstellen van een kansmodel en vervolgens valideren in hoeverre een observatie gegenereerd kan worden door het model. Indien de kans laag is, wordt de observatie gezien als een outlier.

Statistische methoden werken meestal in twee fasen:

1. Training fase: Hierbij wordt een statistisch model bepaald dat zo goed mogelijk past op de data. Dit proces wordt ook wel *model fitting* genoemd.
2. Test fase: Hierbij wordt beoordeeld of een observatie een outlier is, door de kans in het statistische model te bepalen.

Parametrische technieken

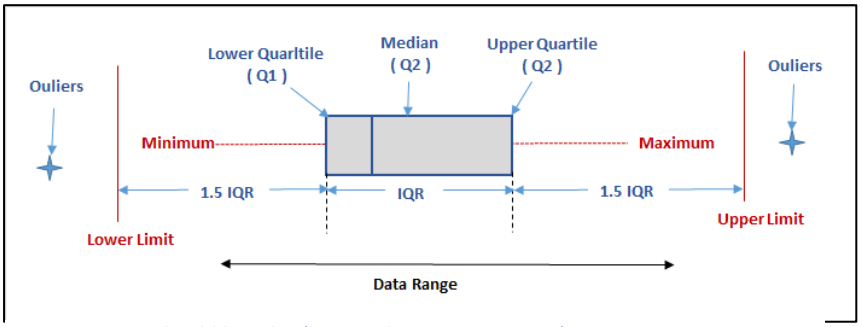
Parametrische technieken gaan ervan uit dat de data gegenereerd is door een bekende distributie, zoals een normale verdeling of poissonverdeling. Op basis van de distributie, kunnen parametrische technieken verder verdeeld worden in:

Gaussiaans model

Dit model gaat uit van een normale verdeling van de data. De training fase bestaat uit het bepalen van de normaal (*mean*) en de variantie (*variance*) m.b.v. MLE (*Maximum Likelihood Estimates*).

Voor de test fase, worden verschillende technieken beschreven, zoals:

- Boxplot rule (Laurikkala, Juhola1, & Kentala., 2000)
- Grubbs test (Grubbs, 1969), (Stefansky, 1972), (Anscombe & Guttman, 1960)
- Rosner test (Rosner, 1983)
- Dixon test (Gibbons, 1994)



Figuur 2.2: Voorbeeld boxplot (bron: WhatIsSixSigma.net)

Regressieanalyse

Outlier detectie door regressieanalyse, wordt vaak toegepast voor tijdreeksen. Hierin wordt onderscheid gemaakt in twee typen outliers.

1. *Waarneembare (Observational) outliers*: Deze treden op wanneer een enkele observatie extreem afwijkt.
2. *Innovational outliers*: Hierbij is er een extreme afwijking van een bepaalde observatie, maar heeft deze afwijking ook invloed op de daaropvolgende observaties.

Bij regressieanalyse, wordt in de training fase een model gezocht dat bij de data past. In de test fase wordt beoordeeld in hoeverre iedere instantie bij het model past.

In diverse technieken (Abraham & Chuang, Outlier detection and time series modeling, 1989), (Abraham & Box, Bayesian analysis of some outlier problems in time series, 1979); (Fox, 1972) wordt de meest aannemelijke schatter (*maximum likelihood estimates)* methode toegepast, waarbij schatting van een parameter die waarde gekozen wordt, waarvoor de aannemelijkheidsfunctie maximaal is.

Niet-parametrische technieken

Outlier detectietechnieken in deze categorie zijn niet gebaseerd op aannames over de distributie van de data. De populairste benaderingen hiervoor zijn Histogram en Finite State Machines.

Histogram

In een histogram wordt geteld hoe vaak een waarde van een feature gemeten wordt, zodat de norm en de verdeling bepaald kan worden. Om te bepalen of een observatie een outlier is, wordt gekeken hoe vaak een bepaalde waarde voorkomt. In de literatuur wordt histogram ook vaak aangeduid als *frequency based* or *counting based*.

Deze techniek wordt veelal semi-supervised toegepast. (Anderson, Frivold, Tamaru, & Valdes, 1994); (Javitz & Valdes, 1991); (Helman & Bhangoo, 1997) gaan ervan uit dat de normale gevallen bekend zijn. (Dasgupta & Nino, 2000) gaat daarentegen uit van bekende outliers. Indien het aantal outliers relatief laag is, kan histogram ook unsupervised toegepast worden.

Histogram gebaseerde technieken wordt veel toegepast in Intrusion Detection (Eskin E. , 2000), (Eskin & Stolfo, 2001) en fraude detectie (Fawcett & Provost, 1999).

Finite State Machines (FSA)

Finite State Machines (Ilgun, Kemmerer, & Porras, 1995); (Salvador & Chan, 2003) en Markov modellen (Smyth, 1994) worden toegepast om outliers in sequentiële data te ontdekken op basis van historische toestandswijzigingen.

2.3.2. Algemene toepasbaarheid van outlier-detectiemethoden

Op basis van de algemene artikelen van (Zhang, 2013) en (Chandola, Banerjee, & Kumar, Anomaly Detection : A Survey, 2009) over outlier detectie en de specifieke artikelen per methode, kunnen de onderstaande methodes en varianten afleiden welke mogelijk toepasbaar zijn voor unsupervised outlier-detectie:

Partitioning based clustering

Methode	Toepasbaarheid	Toelichting
k-means	☐	k-means is niet goed toepasbaar voor outlier detectie omdat k-means alle observaties in clusters indeelt; dus ook de outliers.

Density-based clustering

Methode	Toepasbaarheid	Toelichting
DBSCAN	☑	DBSCAN is uitstekend toepasbaar voor outlier-detectie bij meerdere clusters in een multidimensionale omgeving. DBSCAN is erg gevoelig voor de instelling van de density parameter *Bron:* (Ester, Kriegel, Sander, & Xu, 1996)
DeBaCl	☑	Deze Level Set Tree clustering methode is vergelijkbaar met DBSCAN, maar is beter parameteriseerbaar en herkend subclusters. *Bron:* (Kent, Rinaldo, & Verstynen, 2013)
OPTICS	☑	Deze Level Set Tree clustering methode is vergelijkbaar met DBSCAN en DeBaCl, maar ondersteunt een interactieve analyse en kan clusters van verschillende dichtheid herkennen. *Bron:* (Ankerst, Breunig, Kriegel, & Sander, 1999)

Parametrische methoden

Methode	Toepasbaarheid	Toelichting
Boxplot	☑	Boxplot is toepasbaar voor datasets die uit één cluster bestaan, met een normale verdeling.
RANSAC (Regressie)	☑	Deze vorm van regressieanalyse is toepasbaar voor datasets die uit één cluster bestaan, met afhankelijke variabelen. In tegenstelling tot de reguliere regressieanalyse, houdt RANSAC rekening met outliers bij het bepalen van de regressielijn
Gaussian Mixture Model	☑	Deze vorm van regressieanalyse is toepasbaar voor datasets die uit meerde clusters bestaan, met een normale verdeling.

Niet-parametrische methoden

Methode	Toepasbaarheid	Toelichting
Histogram	☐	Histogram is uitsluitend unsupervised toepasbaar, wanneer het aantal outliers relatief laag is.
Finite State Machines	buiten scope	Deze methode richt zich op sequentiële data en outliers, welke onder type II en III vallen en dus buiten scope van dit onderzoek.

2.3.3. Metrieken om de effectiviteit te vergelijken

Convolutie Matrix

Om de effectiviteit van een techniek te bepalen, wordt onderscheid gemaakt tussen vier scenario's:

- ❑ True positive: fraude is voorspeld en in werkelijkheid fraude;
- ❑ False positive: fraude is voorspeld, maar in werkelijkheid geen fraude;
- ❑ False negative: geen fraude is voorspeld, maar in werkelijkheid wel fraude;
- ❑ True negative: geen fraude is voorspeld en in werkelijkheid geen fraude;

Schematisch kan dit weergegeven worden in een zogenaamd Convolutie Matrix:

		Voorspeld	
		Fraude	**Legitiem**
Werkelijk	**Fraude**	**TP**	**FN**
	Legitiem	**FP**	**TN**

Figuur 2.3: Convolutiematrix

In de ideale situatie, zou een techniek uitsluitend *True Positives* en *True Negatives* op moeten leveren.

Recall ofwel *Sensitivity* geeft de verhouding aan tussen de voorspelde relevante gevallen en de werkelijk relevante gevallen:

Precision

De term *Precision* geeft aan hoeveel van de voorspelde gevallen daadwerkelijk relevant zijn:

$$Precision = \frac{TP}{TP + FP}$$

In een statistische benadering is *Precision* de kans dat een als relevante beoordeelde observatie, daadwerkelijk relevant is.

Recall

Recall ofwel *Sensitivity* geeft de verhouding aan tussen de voorspelde relevante gevallen en de werkelijk relevante gevallen:

$$Recall = \frac{TP}{TP + FN}$$

In een statistische benadering is *Recall* de kans dat een relevante observatie, daadwerkelijk als relevant beoordeeld is.

F1-score

De F1-score ofwel F-score, is de gewogen score van zowel de *recall* als de *precision*:

$$F1\text{-}score = \frac{recall\ x\ precision}{(reacll + precision)}$$

Precision en *recall* worden meestal niet afzonderlijk besproken maar in samenhang met elkaar, zoals een vaste *recall*, een vaste *percision*, F1-score en de Matthews correlatiecoëfficiënt.

Effectiviteit

Om de effectiviteit van een techniek te bepalen, dient er gekeken te worden naar een aantal factoren.

Een *precision* van 1.0 betekent dat ieder voorspeld geval correct is, maar niet of alle relevante positieve gevallen ook daadwerkelijk positief beoordeeld zijn.

Een *recall* van 1.0 daarentegen betekent dat alle relevante gevallen correct voorspeld zijn, maar zegt niets over het aantal gevallen dat incorrect als relevant beoordeeld is. De eenvoudigste manier om een *recall* van 1.0 te bewerkstelligen is om alle gevallen als positief te beoordelen.

2.4. Declaratiefraude scenario's

Het Rapport Onderzoek Zorgfraude (NZa, 2014) van de Nederlandse Zorgautoriteit, heeft de meest voorkomende fraudescenario's in de Zorg onderzocht. In bijlage 1 staan alle scenario's gedetailleerd beschreven.

Hieruit zijn de onderstaande scenario's per zorgsoort uitgekomen:

Zorgsoort	#Scenario's	#Relevant	%Relevant
Geestelijke Gezondheidszorg	15	13	87%
Extramurale Farmaceutische zorg	13	11	85%
Huisartsenzorg	14	13	93%
Mondzorg	12	7	58%
Medisch specialistische zorg	15	11	73%
Fysiotherapie	12	7	58%
Totaal	**81**	**62**	**77%**

Figuur 2.4: Fraude scenario's per zorgsoort

2.4.1. Conclusies

Uit de literatuurstudie blijkt dat er verschillende outlier-detectiemethoden zijn, die mogelijk toepasbaar zijn voor de declaratiefraudescenario's, waarbij er geen gelabelde voorbeelden zijn. Dit zijn de *unsupervised* methoden. De meeste methoden hebben diverse varianten met specifieke eigenschappen.

3. Doel van het empirisch onderzoek

Het empirisch onderzoek, moet leiden tot een keuze-framework, dat gebruikt kan worden om o.b.v. kenmerken van fraudescenario's, de meest geschikte outlier-detectiemethode(n) te kiezen.

Merk op dat het onderzoek zich niet richt op afzonderlijke declaraties, maar op het vinden van zorgaanbieders, die structureel fraude plegen met zorgdeclaraties

Om tot een keuze-framework voor outlier-detectiemethoden voor declaratiefraudescenario's in de Zorg te komen, zijn de onderstaande deelvragen geformuleerd:

5. Wat is een keuze-framework voor outlier-detectiemethoden voor declaratiefraude scenario's in de Zorg?
6. Hoe kunnen declaratiefraude scenario's geclassificeerd worden?
7. Wat is het motivatie voor een keuze van outlier-detectiemethoden voor de klassen declaratiefraude scenario's?

In het empirisch onderzoek moet allereerst onderzocht worden, hoe een keuze-framework eruit moet zien. Op basis van welke kenmerken moeten keuzes gemaakt worden en hoe ziet het resultaat er dan uit dat het keuze-framework levert? Hoe kan het framework in de toekomst ook nieuwe scenario's en nieuwe outlier-detectiemethoden ondersteunen?

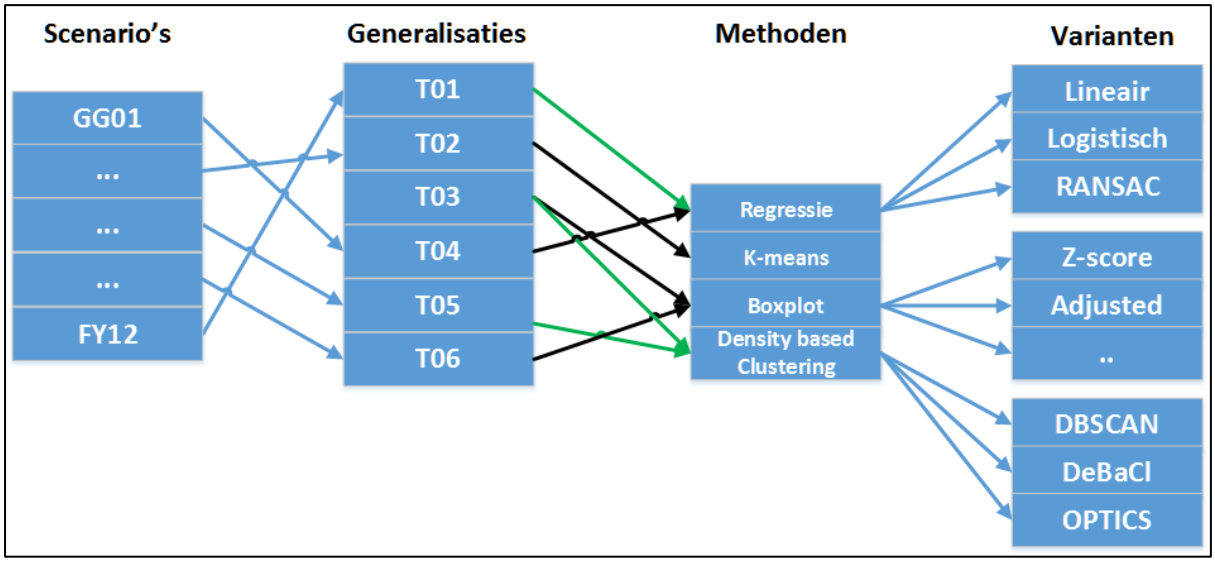
Om het keuze-framework breed inzetbaar te maken, zal er niet uitsluitend gekeken moeten worden naar bekende scenario's. De bekende declaratiefraude scenario's moeten worden geclassificeerd en gegeneraliseerd, zodat het keuze-framework o.b.v. generieke kenmerken ondersteuning kan geven.

Opdat het keuze-framework de meest geschikte outlier-detectiemethode(n) kan bepalen, zullen de verschillen tussen de diverse outlier-detectiemethoden in kaart gebracht moeten worden en de toepasbaarheid op basis van de generieke kenmerken.

4. Methode

4.1. Onderzoeksmodel

De aansluiting tussen de fraudescenario's en de outlier-detectiemethoden, is niet triviaal. Het onderstaande onderzoeksmodel is opgesteld, om gestructureerd tot een mapping te komen, waarbij herleidbaar blijft, hoe deze tot stand gekomen is.



Figuur 4.1: Onderzoeksmodel

Uitgangspunt zijn de bekende fraudescenario's, die in de literatuurstudie naar voren zijn gekomen en in Bijlage 1: Fraude scenario's vermeld zijn. De fraudescenario's worden o.b.v. kenmerken, gegeneraliseerd tot een beperkt aantal generalisaties. Voor iedere generalisatie wordt op basis van de literatuur en testdata, de meest geschikte outlier-dectectiemethoden bepaald. De meeste methoden, kennen verschillende varianten. Een gedetailleerd advies over het gebruik van de varianten, valt buiten de scope van het onderzoek.

4.2. Onderzoeksstrategie

Deelvragen

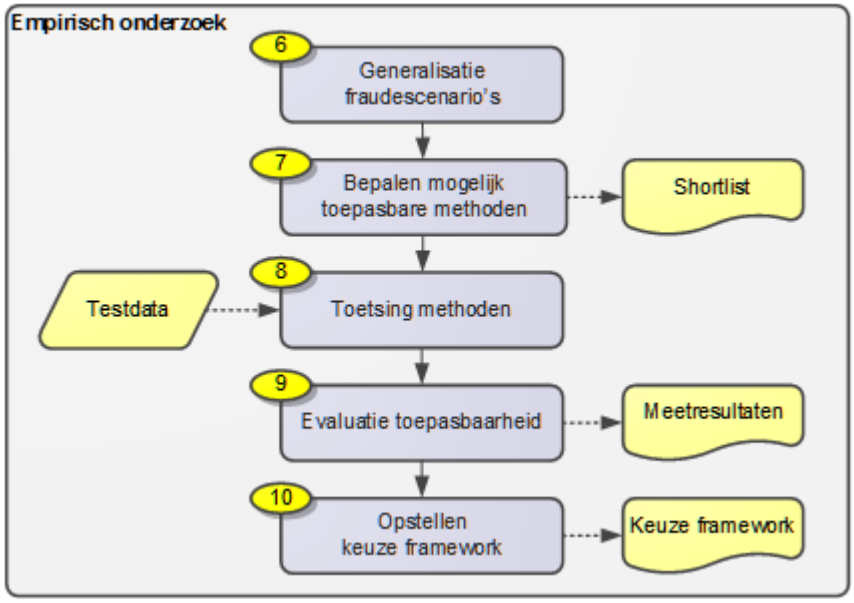
5. Wat is een keuze-framework voor outlier-detectiemethodes voor declaratiefraude scenario's in de Zorg?
6. Hoe kunnen declaratiefraude scenario's geclassificeerd worden?
7. Wat is het motivatie voor een keuze van outlier-detectiemethoden voor de klassen declaratiefraude scenario's?

Vraag 5. Om de classificatiemethode te bepalen, wordt gekeken naar de algemene toepasbaarheid van outlier-detectie methoden op basis van de kenmerkende variabelen.

Vraag 6. Vergelijking van de toepasbare methoden bij iedere klasse op basis van kwaliteit van de outlier detectie.

Vraag 7. Per generalisatie, worden voor alle mogelijk toepasbare outlier-detectiemethoden, de belangrijkste kwaliteitsattributen vergeleken.

Het empirisch onderzoek, wordt uitgevoerd volgens onderstaand stappenplan:



Figuur 4.2: Onderzoeksstrategie

Generalisatie fraudescenario's [stap 6]

In het Rapport Onderzoek Zorgfraude (NZa, 2014), worden de veel voorkomende bekende fraudescenario's, per zorgsoort uiteengezet.

Het te ontwikkelen keuze-framework moet zodanig opgezet worden, dat het niet alleen voor bekende scenario's toepasbaar is, maar ook voor nieuwe en mogelijk nog onbekende scenario's. Het voldoet daarom niet om uitsluitend de beschreven scenario's te bekijken. Daarom is gekozen voor een meer generieke opzet, waarbij gekeken wordt naar de typering van de verschillende scenario's.

Deze typering is bepaald aan de hand van eigenschappen die onderscheidend zijn voor de mogelijke toepasbaarheid per scenario, zoals beschreven in hoofdstuk 2.3.2.

De volgende eigenschappen zijn hierbij van belang:

- ❑ Probleem dimensie (het aantal relevante kenmerken);
- ❑ Aantal clusters van “normale zorgaanbieders”;
- ❑ Distributie van de gegevens.

Bepalen mogelijk toepasbare methoden per generalisatie [stap 7]

In de literatuurstudie is een aantal mogelijk toepasbare outlier-detectiemethoden naar voren gekomen, welke beschreven zijn in hoofdstuk 2.3.2. Deze outlier-detectiemethoden zijn echter niet voor alle generalisaties toepasbaar. In de ze stap, worden de o.b.v. de literatuur de mogelijk toepasbare methoden per generalisatie bepaald.

Toetsing methoden [stap 8]

Per generalisatie worden de mogelijk toepasbare outlier-detectiemethoden vergeleken aan de hand van een gegenereerde testset met declaraties. Het zou de voorkeur hebben om te werken met werkelijke productiedata, maar door de privacygevoelige aard van declaratiegegevens, konden zorgverzekeraars deze data niet beschikbaar stellen. De testdata is opgesteld o.b.v. input van domein experts bij Truston Solutions B.V.

Evaluatie toepasbaarheid [stap 9]

Per categorie uit stap 6, wordt de effectiviteit van de methodes vergeleken op basis van de resultaten uit stap 8. Hierbij wordt met name de *recall* en *precission* vergeleken, maar daarnaast worden ook aspecten als performance en tuning meegenomen.

Bij de evaluatie van de toepasbaarheid, zal rekening worden gehouden met:

- ❑ De *recall*, *precision* en *F1-score* uit de toetsing met testdata;
- ❑ Snelheid (complexiteitsgraad) van het algoritme o.b.v. de literatuur;
- ❑ Eventuele bijkomende bijzonderheden.

De snelheid van het algoritme is erg afhankelijk van de implementatie en de indexering van de test dataset. Daarom is de snelheid niet bepaald op basis van metingen maar op de literatuur.

Opstellen keuze-framework [stap 10]

De laatste stap is het opstellen van het uiteindelijke keuze-framework, waarbij een flowchart gemaakt wordt om o.b.v. kenmerken van fraudescanario’s, de meest geschikte outlier-detectie methode(n) gekozen kunnen worden.

4.2.1. Overwegingen

Omwille van de haalbaarheid en doorlooptijd van onderzoek, zijn er de volgende concessies gedaan:

- ❑ De testset bevat geen ruis. Bij verschillende kenmerken, of combinatie van kenmerken, zullen er in de praktijk veel outliers ontstaan door: legitiem afwijkend gedrag, fouten en niet representatieve gemiddelden bij zorgaanbieders met weinig observaties. Bij de selectie van de mogelijk toepasbare methoden, wordt hier wel rekening mee gehouden. Het weglaten van deze ruis, zal een te positief beeld geven op effectiviteit van de methoden, maar zal naar verwachting slechts beperkte invloed hebben op de vergelijking tussen de methoden.
- ❑ Veel methoden kennen afgeleide methoden of meerdere specifieke implementaties. In het onderzoek worden niet alle afgeleide en vergelijkbare methoden meegenomen. Waar mogelijk, zal er wel een verwijzing naar de alternatieve methode of implementatie gemaakt worden.
- ❑ Bij enkele scenario’s, is het toepassen van één enkele methode niet effectief. Hierbij is mogelijk een combinatie van methoden noodzakelijk. Deze zijn buiten beschouwing gelaten.

- Het framework ondersteunt het opsporen van afwijkend gedrag door zorgaanbieders, dus niet het opsporen van individuele declaraties. Hiervoor zijn immers reeds diverse materiële controles bij zorgverzekeraars ingericht.

4.3. Onderzoeksaanpak

4.3.1. Generalisatie fraudescenario's

In het literatuuronderzoek, zijn 62 relevante fraudescenario's naar voren gekomen, verdeeld over 6 zorgsoorten. In samenwerking met domein experts bij Truston Solutions B.V., worden per fraude scenario de relevante kenmerken beschreven. Daarnaast wordt op basis van ervaring, de verdeling van de data van legitieme - en frauduleuze zorgaanbieders geclassificeerd tot generalisaties. Per relevant fraudescenario, worden één of meerdere generalisaties benoemd: $T_1..T_N$, waarbij N het aantal generalisaties.

De generalisatie, vindt plaats op basis van de volgende eigenschappen:

Eigenschappen	Mogelijkheden
Probleem dimensie (het aantal relevante kenmerken)	• 1 • 2 • >2
Aantal clusters van "normale zorgaanbieders"	• 1 • >1
Distributie van de gegevens	• Normale (Gaussiaanse) verdeling • Uniforme verdeling
Correlatie tussen de kenmerken	• Afhankelijk • Niet afhankelijk

Tabel 4.1: Generalisatie kenmerken

Bovenstaande eigenschappen zijn bepaald op basis van de specifieke toepasbaarheid per outlier-detectiemethode, zoals beschreven in hoofdstuk 2.3.2. Op basis van deze eigenschappen, worden de mogelijk toepasbare outlier detatie-methoden per generalisatie aangegeven.

4.3.2. Test dataset

Om de methoden te toetsen, wordt gebruik gemaakt van test datasets, waarbij per generalisatie een test dataset opgezet wordt van een representatief fraudescenario. Hierdoor worden de grafieken concreter en beter te interpreteren. De dataset is gegenereerd op basis van input van domein experts. Omdat niet vastgesteld kan worden hoe representatief deze gegenereerde data is, kan uit het onderzoek de effectiviteit van een methode ook niet met een redelijke betrouwbaarheid vastgesteld worden. Het onderzoek richt zich echter op het vergelijken van de toepasbaarheid, waarvoor de test dataset wel voldoende representatief moet zijn. Het gaat er hierbij dus niet om dat de test dataset, voldoende representatief is voor de werkelijke productiedata, maar voldoende representatief voor de gehele generalisatie. Voor een uitgebreide motivatie en de bespreking van de beperkingen, zie hoofdstuk 7.

De test dataset bestaat uit 1.000 zorgaanbieders, waarvan 5% (=50) frauduleus is. Bij de generalisaties waarbij sprake is van meerdere clusters, worden er meerdere clusters gegenereerd; het totaalaantal zorgaanbieders blijft echter gelijk.

De test dataset wordt zodanig gegenereerd dat er overlap is tussen legitieme en frauduleuze zorgaanbieders. De kans dat alle frauduleuze zorgaanbieders, gelabeld zullen worden als outlier, zonder ook legitieme zorgaanbieders foutief te labelen, is hiermee bijna uitgesloten. Uitgangspunt is om de overlap minimaal 10% te houden, zodat de methode naar verwachting een maximale f1-score van 0.9 zal behalen.

In bijlage 2 staat de Python Jupyter code, waarmee de test datasets gegenereerd zijn.

4.3.3. Tooling

Er zijn verschillende tools, die het datamining proces ondersteunen, zoals R Studio en Jupyter Notebook. R Studio werkt met de taal R en Jupyter Notebook werkt met Python. Voor de meeste datamining technieken, zijn R en Python libraries geïmplementeerd en vrij te gebruiken.

Voor het onderzoek is gekozen voor Jupyter Notebook, omdat deze omgeving de meest uitgebreide libraries heeft om testdata te genereren, data te manipuleren en analyses te visualiseren.

Daarnaast wordt in Jupyter Notebook, de documentatie, de code en de grafieken overzichtelijk gepresenteerd.

4.3.4. Toetsing

Per generalisatie, worden op basis van de test dataset, de mogelijk toepasbare outlier-detectiemethoden geïmplementeerd en geoptimaliseerd. Om de resultaten goed te kunnen vergelijken, worden de parameters zodanig ingesteld dat de *recall* in alle gevallen hetzelfde is en op basis van de *precision* vergeleken kan worden hoeveel van de gelabelde outliers juist waren. In enkele gevallen, zal het voorkomen dat de *recall* niet eens op een vergelijkbaar niveau te parameteriseren is.

Per generalisatie, wordt voor iedere outlier-detectiemethode een convolutiematrix bepaald:

		Voorspeld	
		Fraude	**Legitiem**
Werkelijk	**Fraude**	**TP**	**FN**
	Legitiem	**FP**	**TN**

Tabel 4.2: Template convolutiematrix

Per generalisatie, wordt op basis van de convolutiematrix, de *recall*, *precision* en *f1-score* berekend, bijvoorbeeld:

Methode	Recall	Precision	F1-score
Methode 1	0.75	0.90	0.82
Methode 2	0.75	0.88	0.81
Methode 3	0.73	0.90	0.80

Tabel 4.3: Template effectiviteit score

4.3.5. Opstellen keuze-framework

Op basis van de kenmerken in stap 6 en de evaluatie van de toepasbaarheid, wordt het keuze-framework uitgewerkt in een beslisboom.

Het keuze-framework voor outlier-detectiemethoden voor declaratiefraude scenario's in de Zorg, is een raamwerk waarmee data-analisten per probleemdomein de meest geschikte methode(n) kunnen bepalen. Het keuze-framework is toepasbaar voor het opsporen van zorgaanbieders, die een afwijkend gedrag vertonen in hun declaraties.

Omdat de aanbevelingen voor de outlier-detectiemethoden gebaseerd zijn op meerdere eigenschappen van het probleemdomein of declaratiefraude scenario, wordt het framework opgezet in de vorm van een flowchart. Hierdoor worden de verschillende keuzemomenten duidelijk weergegeven.

5. Uitvoering

5.1. Het keuze-framework

Een soortgelijk keuze-framework voor outlier-detectiemethoden, is beschreven in het rapport van Songwon Seo (Songwon, 2006). Hierbij wordt het keuze-framework opgesteld aan de hand van een beslisboom, waarbij op kenmerken van het probleemdomein, de meest geschikte methode bepaald wordt. Deze aanpak is ook toepasbaar voor het keuze-framework voor declaratiefraude scenario's in de Zorg.

Het keuze-framework zou te complex worden wanneer alle bekende varianten van de gebruikte outlier-detectiemethoden verwerkt zouden worden. Daarom worden enkele varianten separaat toegelicht.

5.2. Kenmerken

De fraude scenario's, zoals beschreven in bijlage 2, hebben betrekking op declaraties. Het framework dient echter te ondersteunen bij het opsporen van zorgaanbieders die structureel fraude plegen. De classificatie wordt daarom ook op kenmerken op het niveau van de zorgaanbieder gedaan. Dit zijn dus vaak afgeleide en geaggregeerde kenmerken, zoals: aantal patiënten, omzet, omzet/patiënt, #sessies/patiënt.

5.3. Classificatie van declaratiefraude scenario's

Voor de toepasbaarheid van het keuze-framework, is een eenduidige classificatie noodzakelijk. De classificatie wordt gedaan op basis van de volgende aspecten met betrekking tot de kenmerken van zorgaanbieders:

- ❑ Typering van de kenmerken;
- ❑ Correlatie tussen de kenmerken;
- ❑ Distributie van de gegevens.

5.3.1. Typering van kenmerken

Uit de fraudescenario's in bijlage 1, kunnen de volgende typen kenmerken worden afgeleid:

Enkelvoudige kenmerken

De enkelvoudige kenmerken zijn veelal gerelateerd aan de grootte van de zorgaanbieder en zeggen op zichzelf weinig over het mogelijk afwijkend gedrag.
Voorbeelden hiervan zijn:

- ❑ #patiënten;
- ❑ #bezoeken;
- ❑ omzet(€);
- ❑ #behandelingen/jaar.

Samengestelde kenmerken

Samengestelde kenmerken zijn opgebouwd uit twee of meer enkelvoudige kenmerken en zeggen vaak iets over het gedrag, in het bijzonder dus ook over het afwijkend gedrag en mogelijke fraude. Samengestelde kenmerken uitten zich veelal in de vorm van en percentage of gemiddelde.
Voorbeelden hiervan zijn:

- ❑ %onverzekerde zorg = #onverzekerde zorg / (#verzekerde zorg + #onverzekerde zorg)
- ❑ Øomzet per patiënt (= omzet / #patiënten)

Categorische kenmerken

De categorische kenmerken typeren de zorgverlever, de behandeling of de patiënt.
Afwijkend gedrag kan geanalyseerd worden door de afwijkende verdeling van aantallen binnen een categorie. Een bijzonder geval hierbij zijn de dichotome kenmerken, welke slechts 2 mogelijke uitkomsten heeft.
Voorbeelden hiervan zijn:

- ❑ Patiënt: Man, Vrouw
- ❑ Declaratie: ANW-toeslag (ja/nee)
- ❑ Tandarts behandeling: eenvlaksvulling, tweevlaksvulling, kroon, drie-delige brug, ..

5.3.2. Generalisaties

Om afwijkend gedrag te kunnen constateren, wordt getracht het normale gedrag te modelleren, opdat afwijkend gedrag afleidbaar is uit de outliers. Om het normale gedrag te modelleren wordt uitgegaan van de data distributie van de kenmerken.

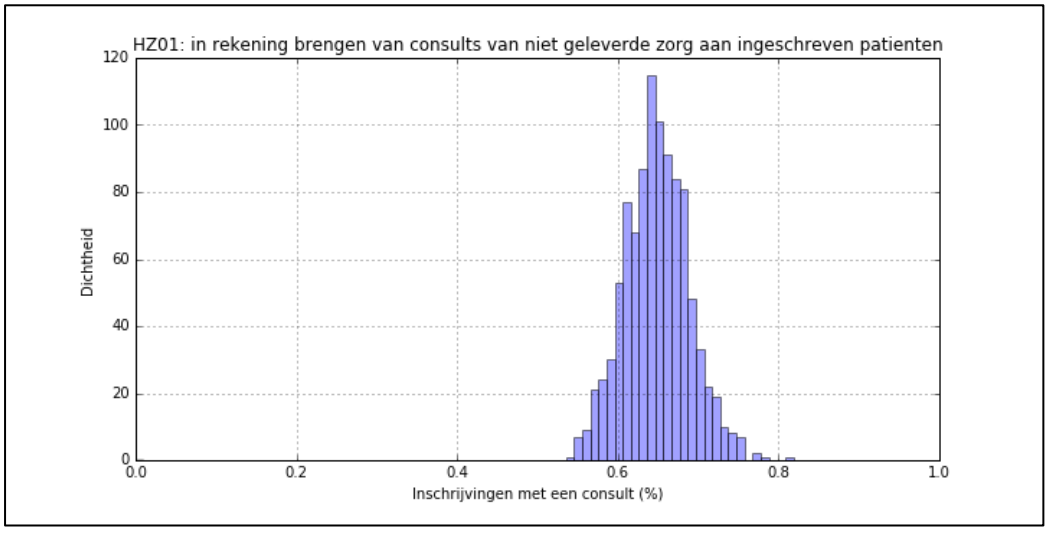
Hierbij kan onderscheid gemaakt worden in de volgende verdelingen:

Generalisatie T0: Uniforme verdeling

Bij een uniforme verdeling, is de kans op alle waardes even groot. Kenmerken met een uniforme verdeling zijn in de beoordeelde scenario's echter niet bruikbaar omdat de kans op alle waarden even groot is. Deze generalisatie wordt dan ook verder buiten beschouwing gelaten.

Generalisatie T1: Normale (Gaussiaanse) verdeling met één variabele (1D)

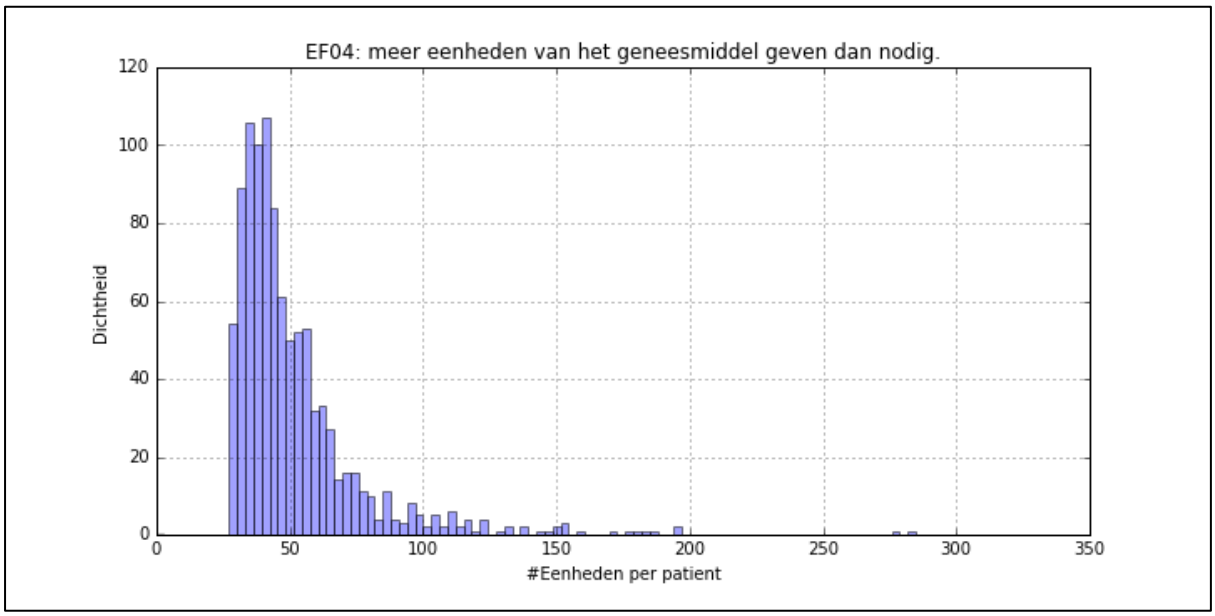
Bij de meeste samengestelde kenmerken, is de kansdichtheid verdeeld volgens een normale verdeling, ook wel Gaussiaanse verdeling genoemd.



Figuur 5.1: Normale verdeling

Generalisatie T2: Lognormale verdeling met één variabele (1D)

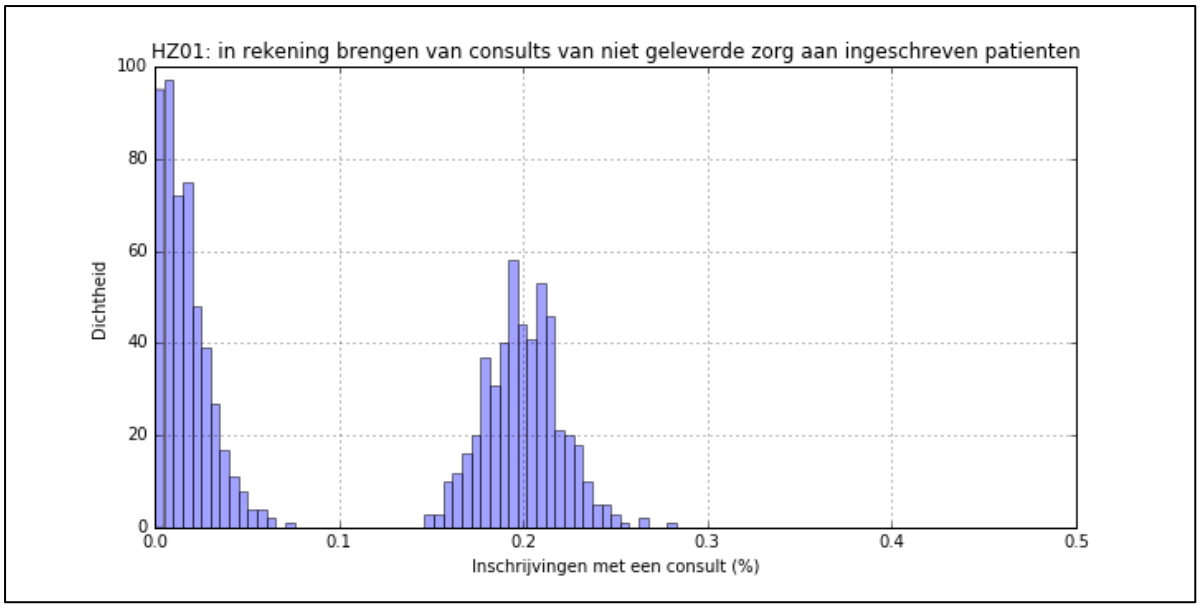
Hoewel de normale verdeling, zoals beschreven in generalisatie 1a, in de wetenschap vaak gebruikt wordt, is de verdeling in de praktijk vaak *skewed*, met name wanneer er sprake is van samengestelde kenmerken. Dit fenomeen wordt uitvoerig beschreven door (Limpert, Stahel, & Abbt, 2001), waarbij ook raakvlakken in het Zorg domein aan bod komen, waaronder: Farmacie, sociale wetenschappen en economie. Generalisatie T2 gaat uit van een lognormale verdeling, die volgens (Limpert, Stahel, & Abbt, 2001) het meeste voorkomt.



Figuur 5.2: Lognormale verdeling

Generalisatie T3: Gecombineerde normale verdeling met één variabele (1D)

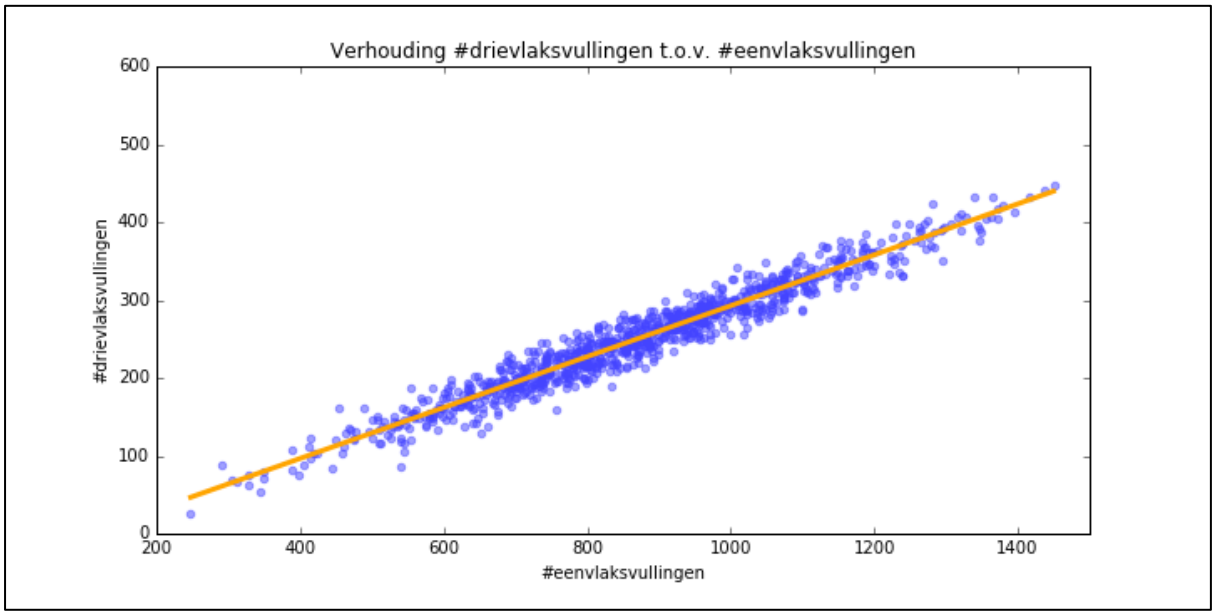
Bij een gecombineerde normale verdeling zijn er meerdere clusters.



Figuur 5.3: Gecombineerde normale verdeling

Generalisatie T4: Twee afhankelijke variabelen (2D)

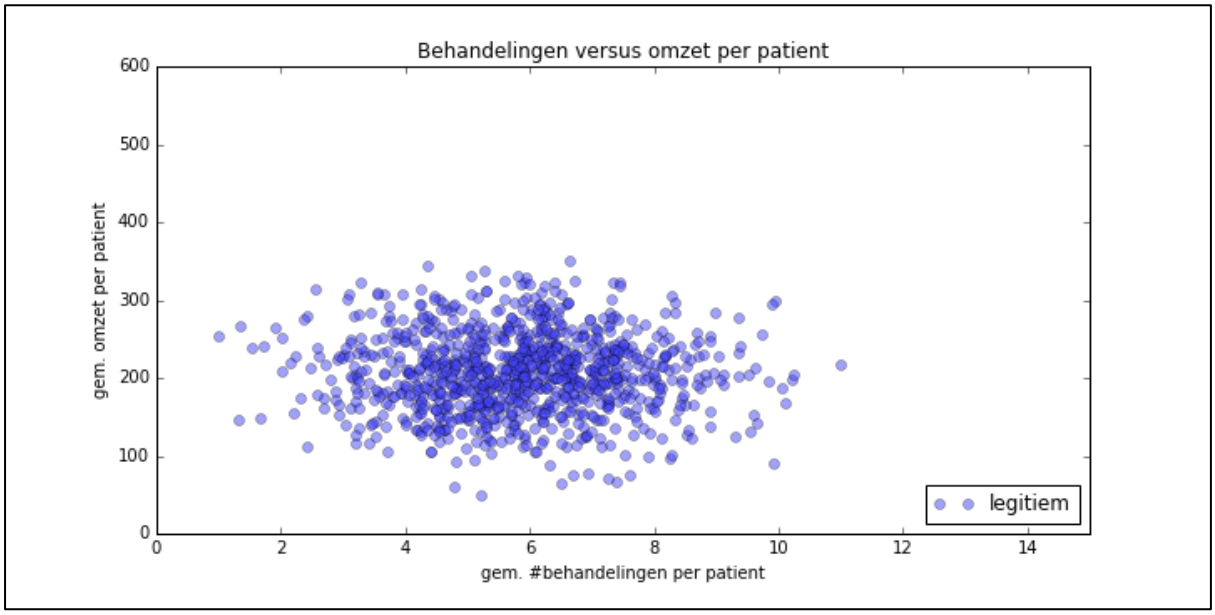
Indien er twee afhankelijke variabelen zijn, kan er een regressielijn getekend worden. Onderstaande figuur toont twee lineair afhankelijke variabelen. Naast een lineair afhankelijkheid, komt een logistische ook veel voor.



Figuur 5.4:Twee afhankelijke variabelen

Generalisatie T5: Eén cluster met twee onafhankelijke variabelen (2D)

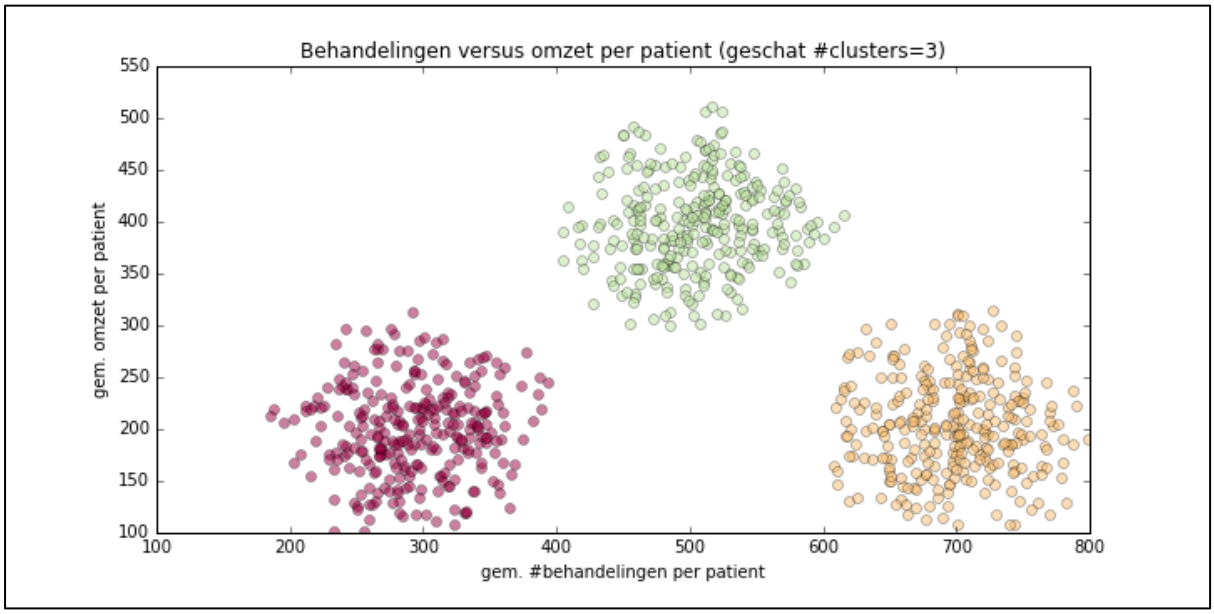
Onderstaande figuur toont een cluster met twee onafhankelijke variabelen.



Figuur 5.5: Cluster met twee onafhankelijke variabelen

Generalisatie T6: Meerdere clusters met twee onafhankelijke variabelen (2D)

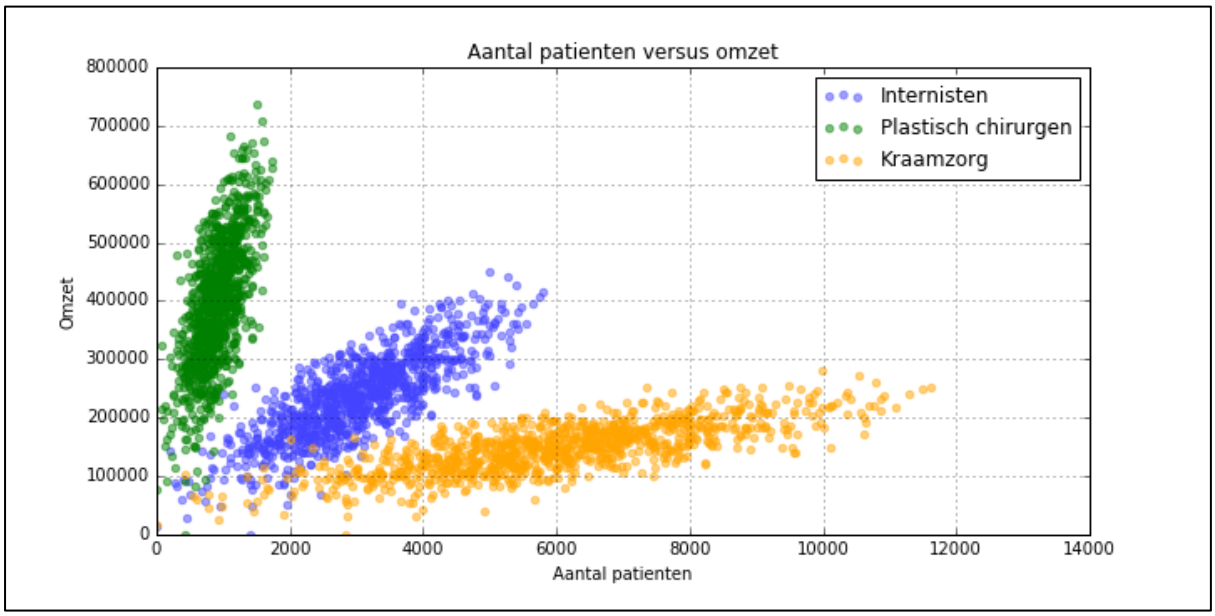
Onderstaande figuur toont drie clusters van twee onafhankelijke variabelen.



Figuur 5.6: Meerdere clusters met onafhankelijke variabelen

Generalisatie T7: Meerdere categorieën van afhankelijke variabelen

Onderstaande figuur toont drie clusters van twee afhankelijke variabelen. Omdat het effectief vinden van outlier, niet met behulp van één outlier-detectiemethode mogelijk is, wordt deze generalisatie niet in verder onderzocht uitgewerkt.



Figuur 5.7: Meerdere categorieën van afhankelijke variabelen

5.4. Motivatie voor scenario's

5.4.1. Selectie van mogelijk toepasbare methoden

De toepasbaarheid van outlier-detectie methoden, is erg afhankelijk van het specifieke probleemdomein. Om tot een shortlist van mogelijk toepasbare methoden te komen, worden de onderstaande criteria beoordeeld. De criteria zijn opgesteld aan de hand van de specifieke kenmerken van de outlier-detectiemethoden in hoofdstuk 2.3.2. De motivatie voor de criteria wordt verderop toegelicht.

Criteria	Eis
Unsupervised	☑
Ongevoeligheid voor outliers	☑
Interactive analyse	☐
Ondersteuning van meerdere clusters	☐ / ☑
Variabele dichtheid	☐
Computationele complexiteit maximaal O(n log n)	☐

Tabel 5.1: Criteria shortlist

Unsupervised, zonder voorkennis

Volgens de NZa (NZa, 2014) en domein experts, hebben zorgverzekeraars onvoldoende gegevens om de fraudegevallen binnen de bestaande dataset te kunnen labelen. Outlier detectie zonder het labelen van data valt onder unsupervised leren. Supervised en semi-supervised methoden zijn dus niet toepasbaar.

Ongevoeligheid voor outliers (zoals noise)

Indien een methode te gevoelig is voor outliers, worden de outliers als de norm beschouwd. Bij k-means clustering worden bijvoorbeeld alle observaties in een cluster opgenomen, dus ook de outliers. Dit geeft een vertekend beeld van het "normale" gedrag.

Interactive analyse

Bij alle outlier-detectie methoden, zijn er parameters waarmee de gevoeligheid instelbaar is. De optimale parameters zijn niet alleen afhankelijk van de verdeling van de data, maar ook van de afweging voor meer potentiele fraudegevallen versus een hogere betrouwbaarheid.

Sommige methodes ondersteunen een interactieve analyse, waarbij het resultaat van parameteraanpassingen direct zichtbaar zijn. Dit criterium is geen harde eis omdat het geen directe invloed heeft op het resultaat van het proces. Er is ook een alternatief waarbij de parametertuning uitgevoerd wordt o.b.v. een kleinere set van zorgaanbieders.

Meerdere clusters

In veel gevallen zullen de kenmerken meerdere clusters bevatten, omdat er ook verschillende type zorgaanbieders zijn. Zo zullen bijvoorbeeld sport-fysiotherapeuten gemiddeld minder kinderfysio behandelingen uitvoeren dan fysiotherapeuten, die gespecialiseerd zijn in fysiotherapie bij kinderen. Uiteraard een eis indien de er meerdere clusters zijn.

Variabele dichtheid

Clusters kunnen een verschillende dichtheid hebben. Een observatie zal bij een cluster met een hoge dichtheid, eerder als outlier gezien moeten worden dan bij een cluster met een lage dichtheid, indien de afstand tot beide clusters gelijk is.

Computationele complexiteit

Zorgverzekeraars verwerken jaarlijks 1,1 miljard (bron ZN https://www.zn.nl/350584836/Feiten-en-cijfers) declaraties van ruim 300.000 (bron BIG register https://www.bigregister.nl/overbigregister/cijfers/) zorgaanbieders. Deze aantallen vereisen een

hoge mate van efficiency om clusters en outliers te bepalen. Een kwadratisch complexiteit ofwel $O(n^2)$, zou voor 300.000 zorgaanbieders een factor van $9*10^{10}$ rekenstappen betekenen. Met de huidige processorkracht zouden deze methodes uitsluitend toepasbaar zijn voor een beperkte set van Zorgaanbieders. Een complexiteit van O(n log n) zou bij 300.000 zorgaanbieders factor van $5,4*10^6$ rekenstappen betekenen, wat tegenwoordig voor een moderne snel te berekenen is.

5.4.2. Kenmerken per outlier-detectiemethode

Op basis van de kenmerken, die in de literatuurstudie per outlier-detectiemethoden naar voren zijn gekomen, kunnen de methoden als volgt geclassificeerd worden:

Methode	Unsupervised	Outliers (ook ruis)	Interactief	Meerdere clusters	Variabele dichtheid	Complexiteit
RANSAC (Lineaire) regressie *Bron:* (Fischler & Bolles, 1981)	+	+	-	-	-	+
GMM (Gaussian Mixture Model) *Bron:* (Aitkin & Wilson, 1980)	+	+	-	+	-	+
Boxplot *Bron:* (Laurikkala, Juhola1, & Kentala., 2000)	+	+	-/+	-	-	+
k-means *Bron:* (Jain K. , 2010)	+	--	-	+	-	-
DBSCAN *Bron:* (Ester, Kriegel, Sander, & Xu, 1996)	+	+	-	+	-	+*
DeBaCl *Bron:* (Kent, Rinaldo, & Verstynen, 2013)	+	+	+	+	+	+*
OPTICS *Bron:* (Ankerst, Breunig, Kriegel, & Sander, 1999)	+	+	+	+	+	+*

Tabel 5.2: Classificatie outlier-detectiemethoden

*: Complexiteit = O(n log n) indien kNN geïndexeerd.

K-means classificeert alle observaties, dus ook de outliers. Daarom wordt k-means verder niet in het onderzoek meegenomen. OPTICS wordt niet verder onderzocht omdat de werking nauw overeenkomt met DeBaCl.

5.5. Toepasbaarheid per generalisatie

Op basis van de kenmerken per generalisatie, zoals deze in hoofdstuk 5.3.2 beschreven zijn, kunnen we de onderstaande tabel afleiden waarbij voor iedere generalisatie, de mogelijk toepasbare outlier-detectiemethoden vermeld staat. T7 wordt buiten beschouwing gelaten omdat deze generalisatie hybride outlier-detectiemethoden vereist.

Generalisatie		RANSAC	GMM	Boxplot	DBSCAN	DeBaCl	OPTICS	[Hybride methodes]
T1	Normale verdeling met één kenmerk	-	•	•	-	-	-	-
T2	Lognormale verdeling met één kenmerk	-	•	•	-	-	-	-
T3	Gecombineerde normale verdeling met één kenmerk	-	•	-	•	•	•	-
T4	Meerdere afhankelijke kenmerken (of categorieën)	•	-	•	-	-	-	-
T5	Eén cluster met twee onafhankelijke kenmerken	-	-	•	•	•	•	-
T6	Meerdere clusters met twee onafhankelijke kenmerken	-	-	-	•	•	•	-
T7	Meerdere klassen van afhankelijke kenmerken	-	-	-	-	-	-	•

Tabel 5.3: Toepasbaarheid outlier-detectiemethoden

Op basis van de werkelijke dataverdeling, kan bepaald worden of een kenmerk onder T1, T2 of T3 valt.

5.6. Empirische toetsing

Voor iedere typering, worden de mogelijk toepasbare outlier-detectie methoden getoetst op basis van een specifieke test dataset.

De test dataset bevat altijd 1000 zorgaanbieders, waarvan 50 (=5%) frauduleus zijn. De test datasets zijn zo opgezet, dat de frauduleuze zorgaanbieders overlappen met de niet-frauduleuze zorgaanbieders. Hierdoor is een F1-score van precies 1.0 uitgesloten.

5.6.1. Generalisatie T1 - Normale verdeling met één kenmerk

Generalisatie gaat T1 gaat uit van de volgende kenmerken:

Kenmerken	Waarde
Probleem dimensie	1
Aantal clusters	1
Distributie	Normale verdeling
Correlatie tussen de kenmerken	n.v.t.

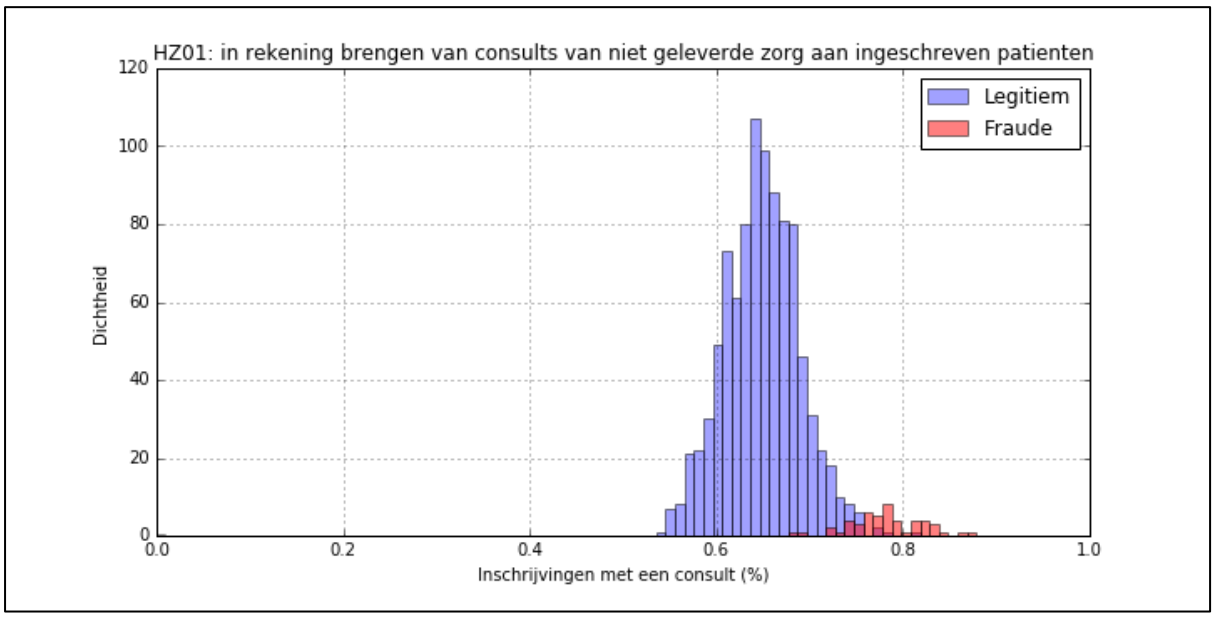
Tabel 5.4: Kenmerken generalisatie T1

Als representant wordt een fraudescenario bij huisartsen genomen, HZ01: in rekening brengen van consults van niet geleverde zorg aan ingeschreven patiënten. Gemiddeld komen 65% van de ingeschreven patiënten jaarlijks minstens één keer bij de huisarts voor een consult, met een spreiding van 4%.

Bij scenario HZ01 zijn er huisartsen die consults bij de zorgverzekeraars declareren van patiënten die dat jaar helemaal niet op de praktijk geweest zijn. Dit zal leiden tot een relaties hoger aantal inschrijvingen met een consult.

Test dataset

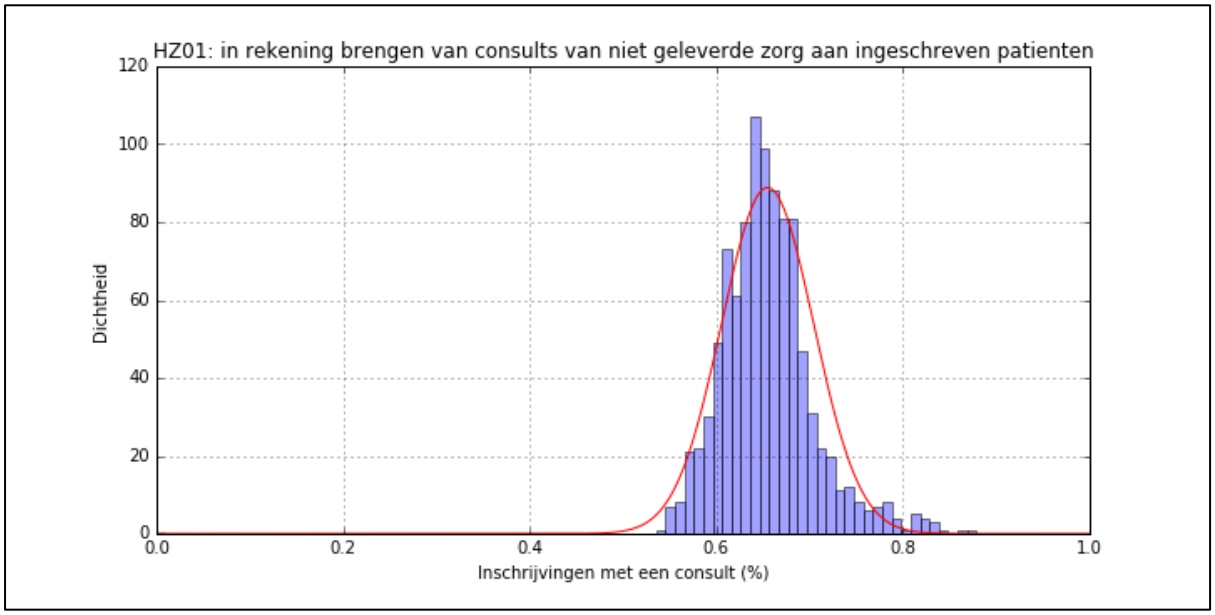
De test dataset hiervoor is als volgt opgebouwd:



Figuur 5.8: Test dataset voor generalisatie T1

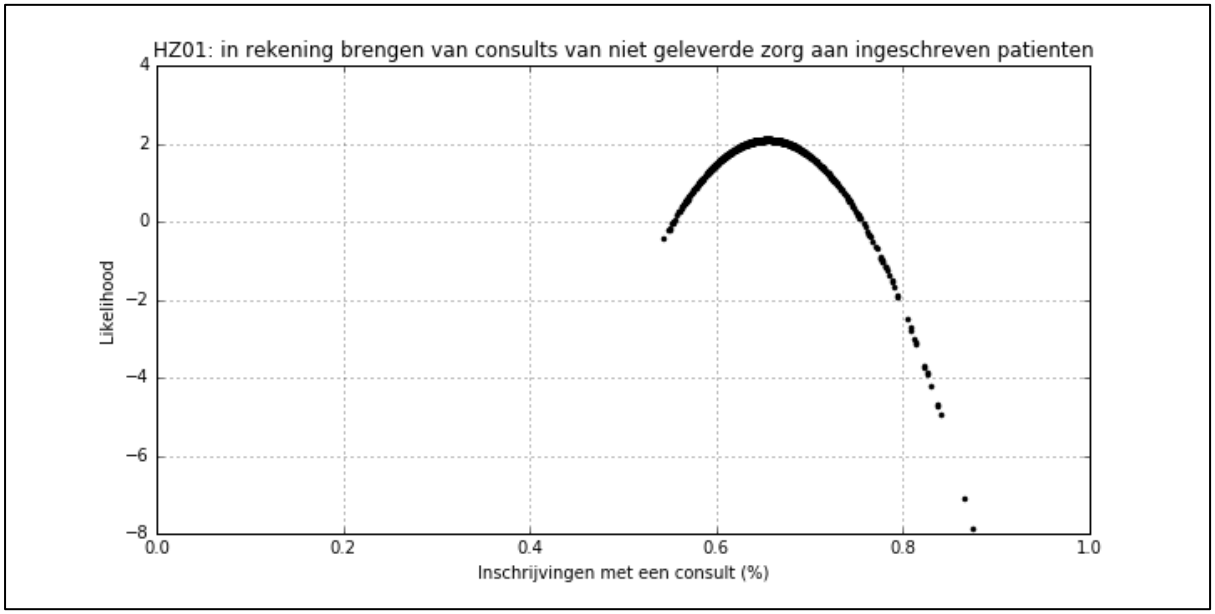
Methode T1a: Gaussian Mixture Model (GMM)

Het Gaussian Mixture Model, probeert bij benadering een normale verdeling te bepalen, die zoveel mogelijk de observaties in de test dataset benadert. In de onderstaande figuur is deze verdeling als rode lijn weergegeven.



Figuur 5.9: GMM voor generalisatie T1

Nu kan voor iedere observatie, bepaald worden hoe groot de kans (likelihood) volgens deze verdeling is:



Figuur 5.10: Likelihood GMM voor generalisatie T1

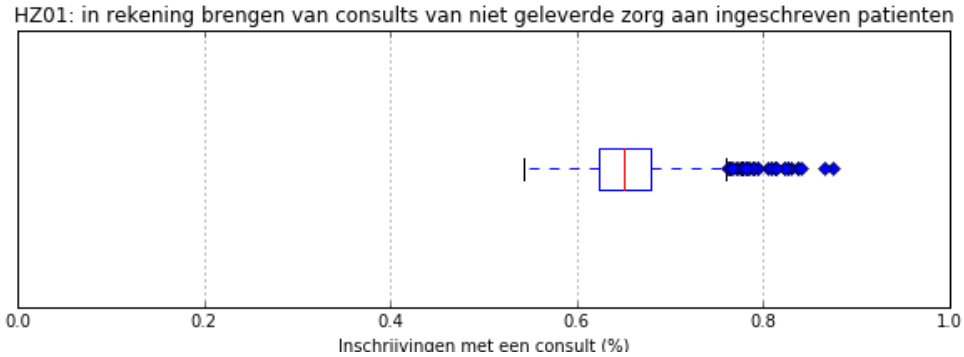
Indien we observaties, als outlier classificeren bij een likelihood < -0.2, levert dit de volgende convolutiematrix:

		Voorspeld	
		Fraude	**Legitiem**
Werkelijk	**Fraude**	36	12
	Legitiem	5	947

Tabel 5.5: Convolutiematrix GMM voor generalisatie T1

Methode T1b: Standaard boxplot

De standaard boxplot hanteert een 1,5 × IKA (interkwartielafstand) en geeft zonder optimalisatie van parameters, het onderstaande resultaat:



Figuur 5.11: Resultaat Boxplot voor generalisatie T1

De convolutiematrix:

		Voorspeld	
		Fraude	**Legitiem**
Werkelijk	**Fraude**	35	13
	Legitiem	4	948

Tabel 5.6: Convolutiematrix standaard boxplot voor generalisatie T1

Methode T1c: Geoptimaliseerde boxplot

Wanneer de parameters optimaliseren naar 1,48 × IKA (interkwartielafstand) wordt het resultaat nog beter:

		Voorspeld	
		Fraude	**Legitiem**
Werkelijk	**Fraude**	36	12
	Legitiem	4	948

Tabel 5.7: Convolutiematrix geoptimaliseerde boxplot voor generalisatie T1

Vergelijking

Methode	Recall	Precision	F1-score
Geoptimaliseerde Boxplot	0.75	0.90	**0.82**
GMM	0.75	0.88	**0.81**
Standaard Boxplot	0.73	0.90	**0.80**

Tabel 5.8: Resultaten generalisatie T1

5.6.2. Generalisatie T2 - Lognormale verdeling met één kenmerk

Generalisatie gaat T2 gaat uit van de volgende kenmerken:

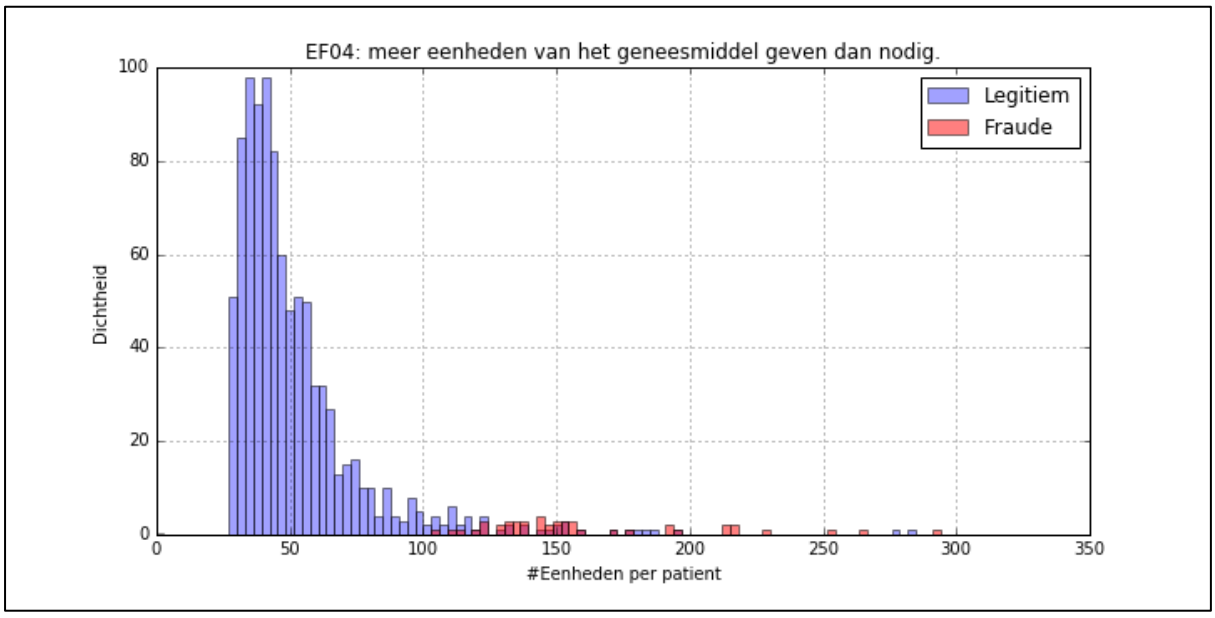
Kenmerken	Waarde
Probleem dimensie	1
Aantal clusters	1
Distributie	Lognormale verdeling
Correlatie tussen de kenmerken	n.v.t.

Tabel 5.9: Kenmerken generalisatie T2

Als representant wordt een fraudescenario bij Farmacie genomen, EF04: meer eenheden van het geneesmiddel geven dan nodig.

Test dataset

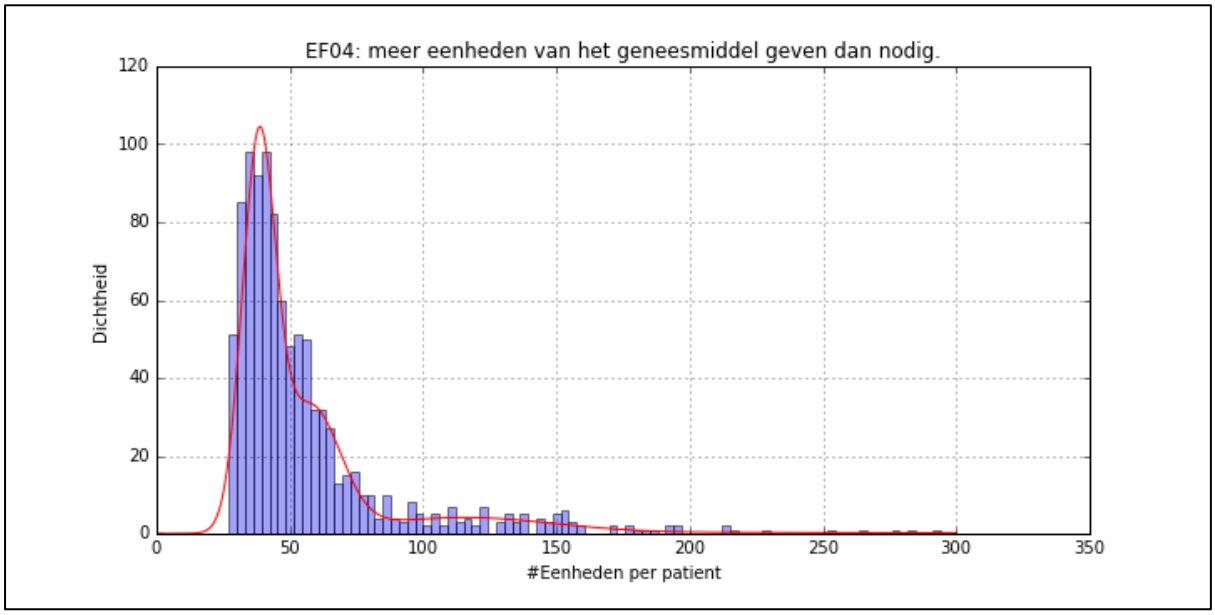
De test dataset hiervoor is als volgt opgebouwd:



Figuur 5.12: Test dataset voor generalisatie T2

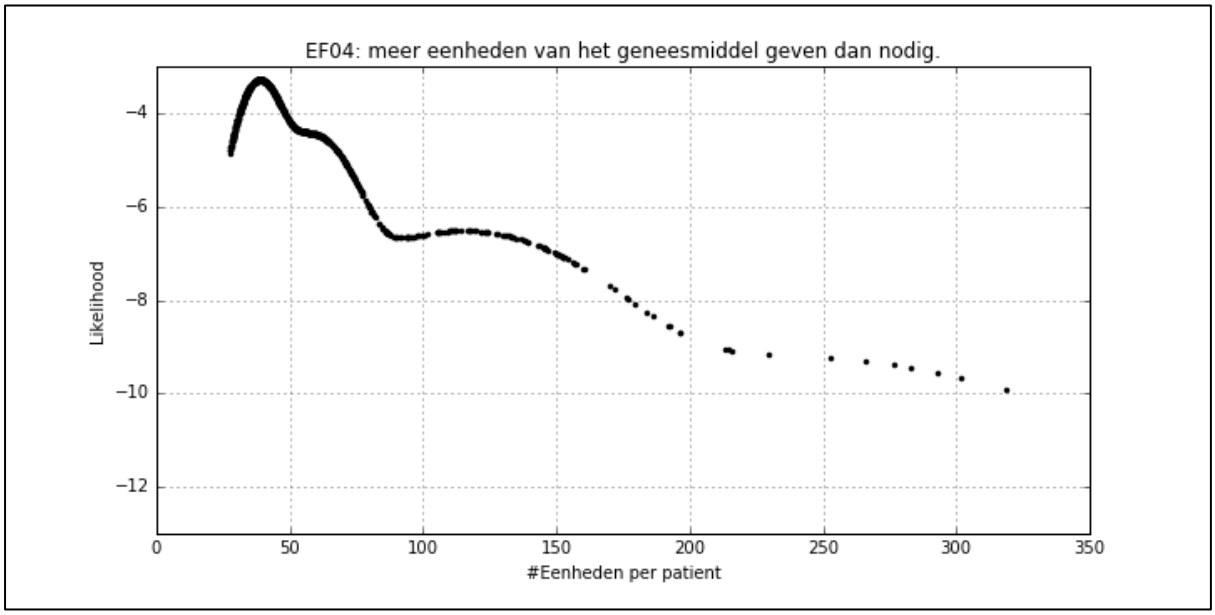
Methode T2a: Gaussian Mixture Model (GMM)

Het Gaussian Mixture Model, probeert bij benadering een normale verdeling te bepalen, die zoveel mogelijk de observaties in de test dataset benadert. In de onderstaande figuur is deze verdeling als rode lijn weergegeven.



Figuur 5.13: GMM voor generalisatie T2

Nu kan voor iedere observatie, bepaald worden hoe groot de kans (likelihood) volgens deze verdeling is:



Figuur 5.14: Likelihood GMM voor generalisatie T2

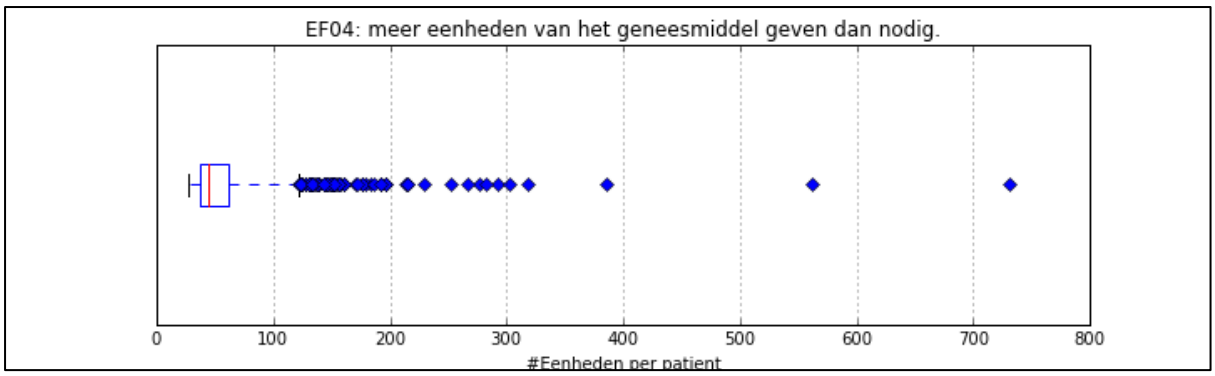
Indien we observaties, als outlier classificeren bij een likelihood < -6.7, levert dit de volgende convolutiematrix:

		Voorspeld	
		Fraude	**Legitiem**
Werkelijk	**Fraude**	**33**	**15**
	Legitiem	**20**	**932**

Tabel 5.10: Convolutiematrix GMM voor generalisatie T2

Methode T2b: Standaard boxplot

De standaard boxplot hanteert een 1,5 × IKA (interkwartielafstand) en geeft zonder optimalisatie van parameters, het onderstaande resultaat:



Figuur 5.15: Resultaat Boxplot voor generalisatie T2

De convolutiematrix:

		Voorspeld	
		Fraude	**Legitiem**
Werkelijk	**Fraude**	**48**	**0**
	Legitiem	**53**	**899**

Tabel 5.11: Convolutiematrix standaard boxplot voor generalisatie T2

Methode T2c: Geoptimaliseerde boxplot

Wanneer de parameters optimaliseren naar 1,48 × IKA (interkwartielafstand) wordt het resultaat nog beter:

		Voorspeld	
		Fraude	**Legitiem**
Werkelijk	**Fraude**	**44**	**4**
	Legitiem	**26**	**926**

Tabel 5.12: Convolutiematrix geoptimaliseerde boxplot voor generalisatie T2

Vergelijking

Methode	Recall	Precision	F1-score
Geoptimaliseerde Boxplot	0.92	0.63	**0.75**
GMM	0.69	0.62	**0.65**
Standaard Boxplot	1.00	0.48	**0.64**

Tabel 5.13: Resultaten generalisatie T2

5.6.3. Generalisatie T3 - Gecombineerde normale verdeling met één kenmerk

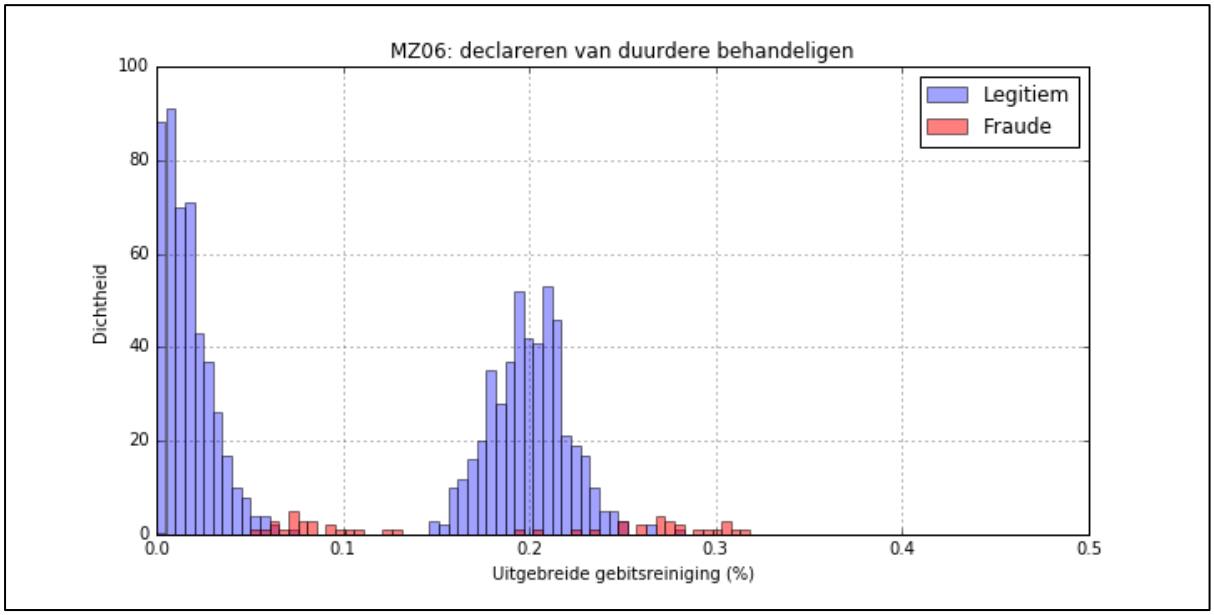
Generalisatie gaat T3 gaat uit van de volgende kenmerken:

Kenmerken	Waarde
Probleem dimensie	1
Aantal clusters	>1
Distributie	Normale verdeling
Correlatie tussen de kenmerken	n.v.t.

Tabel 5.14: Kenmerken generalisatie T3

Een voorbeeld van deze generalisatie is de uitgebreide gebitsreiniging (verwijderen van tandsteen) bij tandartsen. Er zijn tandartsen die gebitsreinigingen zelf uitvoeren, maar veel tandartsen werken hiervoor ook samen met een mondhygiënist. In een histogram, komen deze twee clusters duidelijk naar voren. Een vorm van fraude hierbij is dat tandartsen, naast een reguliere controle, ook een uitgebreide gebitsreiniging declareren, terwijl deze niet heeft plaatsevonden.

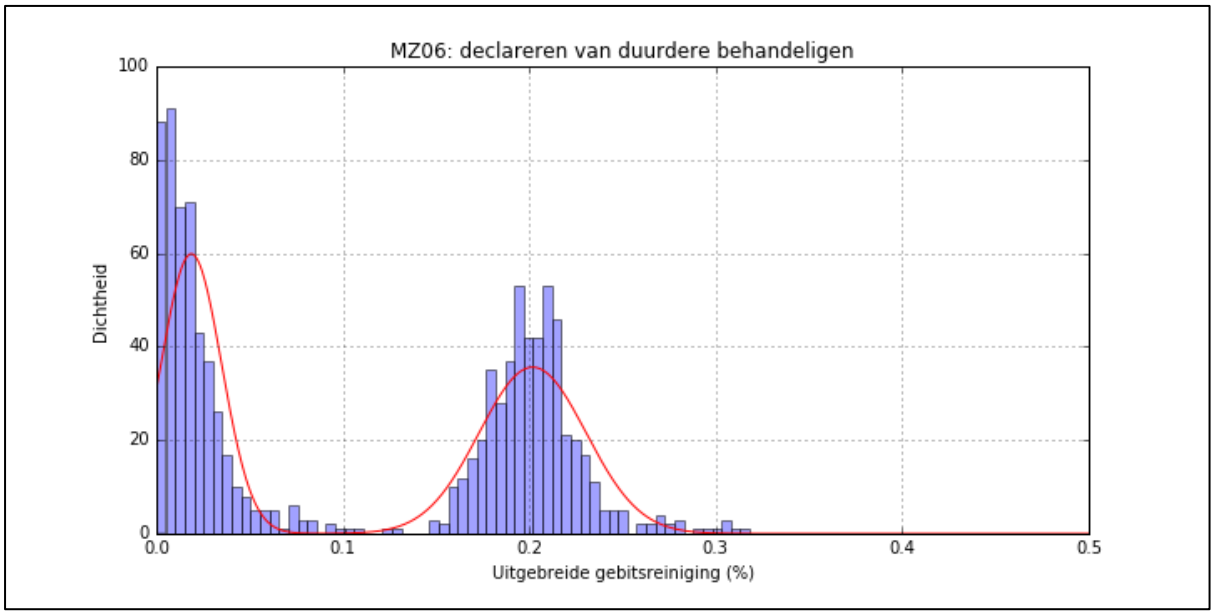
Test dataset



Figuur 5.16: Test dataset voor generalisatie T3

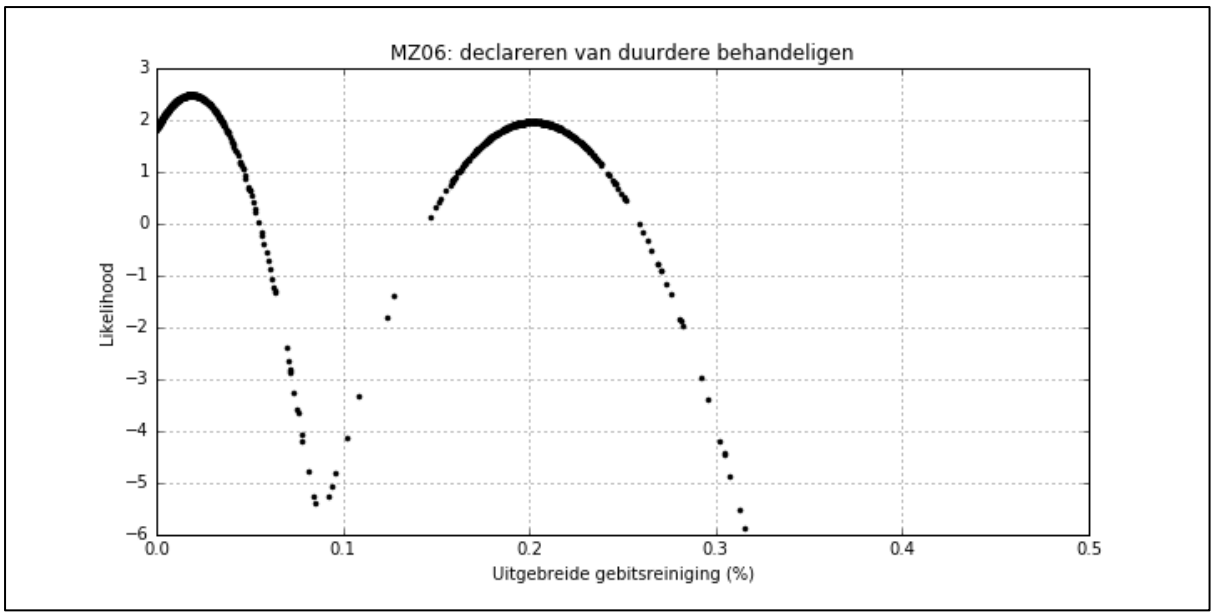
Methode T3a: Gaussian Mixture Model (GMM)

Het Gaussian Mixture Model, probeert bij benadering een normale verdeling te bepalen, die zoveel mogelijk de observaties in de test dataset benadert. In de onderstaande figuur is deze verdeling als rode lijn weergegeven. Een belangrijke parameter bij GMM is het aantal componenten (gaussians) dat gebruikt wordt om de observaties te benaderen. Wanneer



Figuur 5.17: GMM voor generalisatie T3

Nu kan voor iedere observatie, bepaald worden hoe groot de kans (likelihood) volgens deze verdeling is:



Figuur 5.18: Likelihood GMM voor generalisatie T3

Indien we observaties, als outlier classificeren bij een likelihood < 0, levert dit de volgende convolutiematrix:

		Voorspeld	
		Fraude	**Legitiem**
Werkelijk	**Fraude**	**41**	**7**
	Legitiem	**10**	**942**

Tabel 5.15: Convolutiematrix GMM voor generalisatie T3

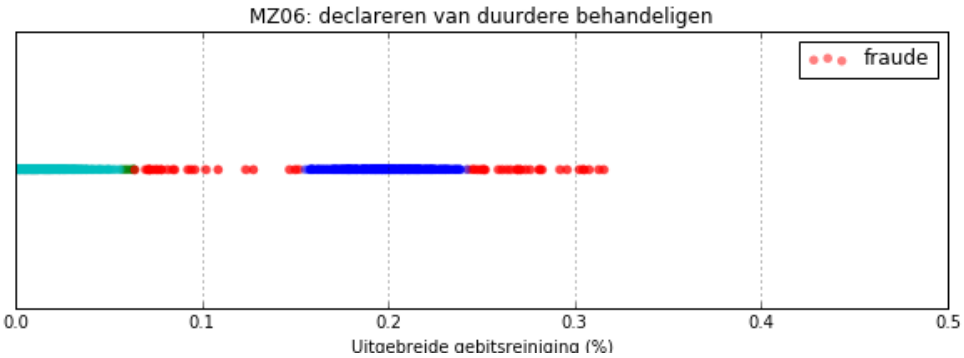
Methode T3b: DBSCAN

DBSCAN is een density-based clustering methode, die meestal toegepast wordt bij meerdere dimensies.

Parameters

ε = 0.004
minPoints = 10

Op de test dataset, vindt DBSCAN de twee clusters terug labelt de observaties die niet tot een cluster behoren als outlier.



Figuur 5.19: Resultaat boxplot voor generalisatie T3

Convolutiematrix:

		Voorspeld	
		Fraude	**Legitiem**
Werkelijk	**Fraude**	**39**	**9**
	Legitiem	**16**	**936**

Tabel 5.16: Convolutiematrix DBSCAN voor generalisatie T3

Methode T3c: DeBaCl

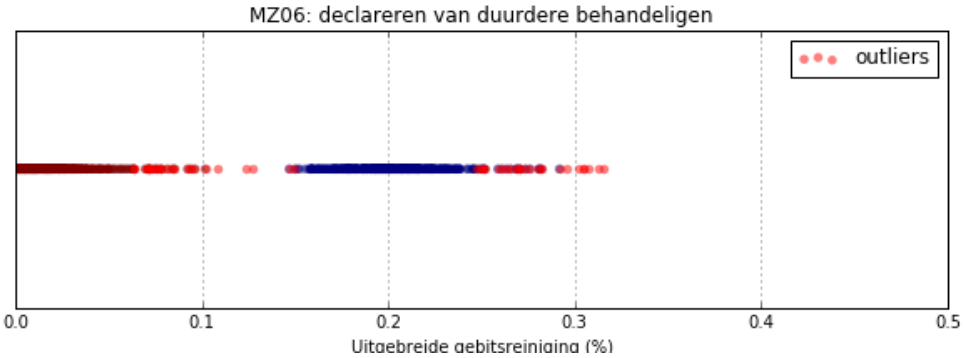
Zoals DBSCAN, wordt de methode DeBaCl ook meestal toegepast bij meerdere dimensies en is minder gevoelig voor parameterinstellingen als DBSCAN.

Parameters

k = 20

Prune threshold = 400

Op de test dataset, vindt DBSCAN de twee clusters terug labelt de observaties die niet tot een cluster behoren als outlier.



Convolutiematrix:

		Voorspeld	
		Fraude	**Legitiem**
Werkelijk	**Fraude**	40	8
	Legitiem	10	942

Tabel 5.17: Convolutiematrix DeBaCl voor generalisatie T3

Vergelijking

Methode	Recall	Precision	F1-score
GMM	0.85	0.80	**0.83**
DeBaCl	0.83	0.80	**0.82**
DBSCAN	0.81	0.71	**0.76**

Tabel 5.18: Resultaten generalisatie T3

5.6.4. Generalisatie T4 - Meerdere afhankelijke kenmerken (of categorieën)

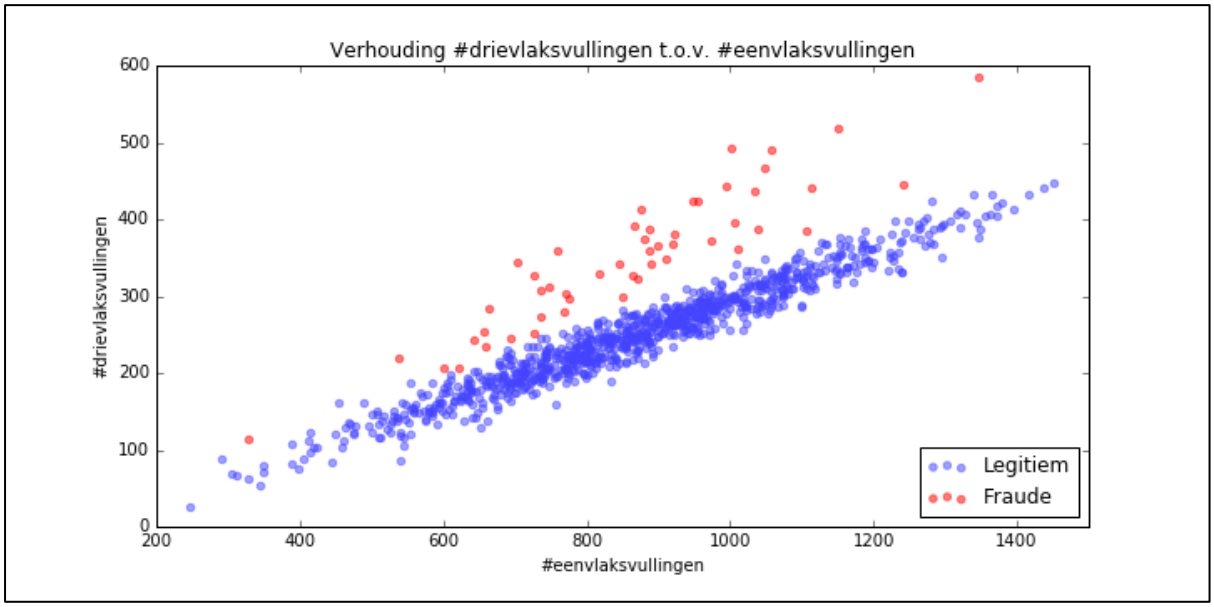
Generalisatie gaat T4 gaat uit van de volgende kenmerken:

Kenmerken	Waarde
Probleem dimensie	>1
Aantal clusters	1
Distributie	N.v.t.
Correlatie tussen de kenmerken	Afhankelijk

Tabel 5.19: Kenmerken generalisatie T4

Een voorbeeld fraudescenario is MZ07, waarbij een tandarts een eenvlaksvulling declareert als een duurdere drievlaksvulling. Uitgangspunt hierbij is dat een zorgaanbieder gemiddeld 3x zoveel eenvlaksvullingen plaatst als drievlaksvullingen.

Test dataset

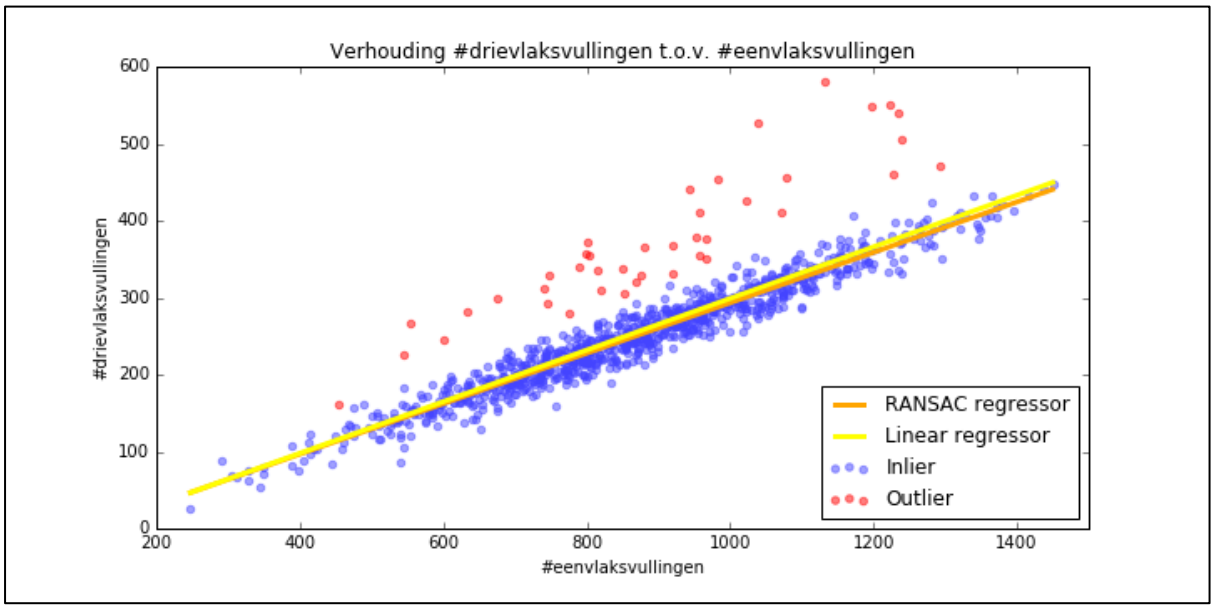


Figuur 5.20: Test dataset voor generalisatie T4

Methode T4a: Lineaire regressie (RANSAC)

Met behulp van lineaire regressie, kan een regressielijn door de observaties getrokken worden. Het nadeel van lineaire regressie, is dat de outliers ook meegenomen worden om het normaal te bepalen. De RANSAC methode, is minder gevoelig voor outliers.

In onderstaande grafiek, zien we het verschil tussen lineaire regressie en RANSAC. Outliers worden op basis van de relatieve afstand tot de regressielijn.



Figuur 5.21: Resultaat RANSAC voor generalisatie T4

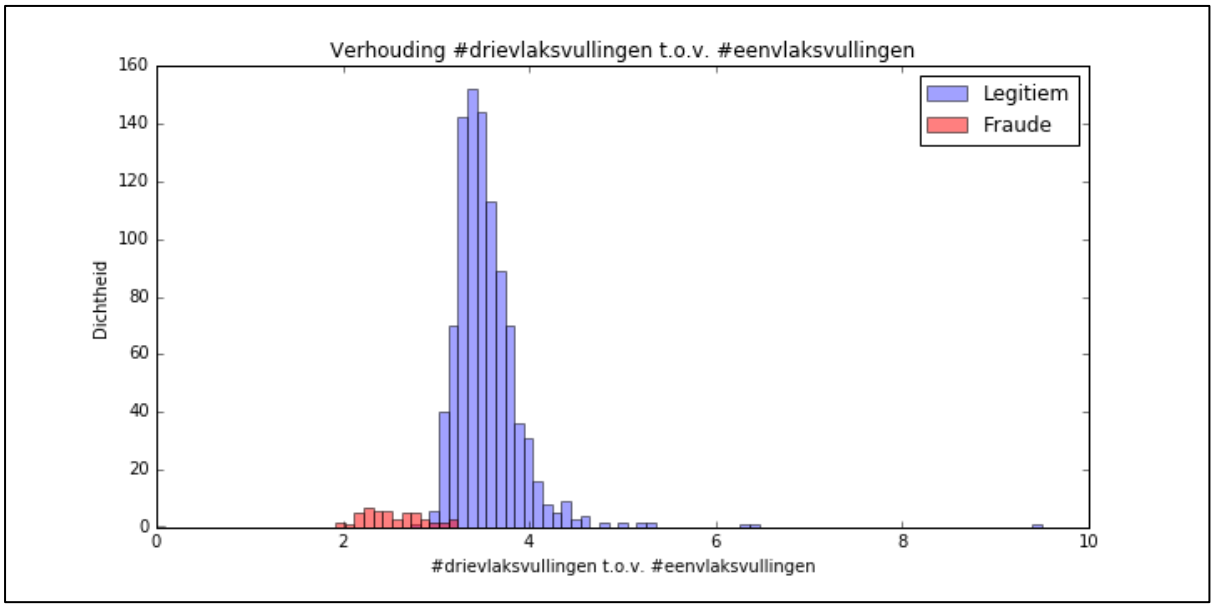
Convolutiematrix:

		Voorspeld	
		Fraude	**Legitiem**
Werkelijk	**Fraude**	**40**	**10**
	Legitiem	**1**	**949**

Tabel 5.20: Convolutiematrix RANSAC voor generalisatie T4

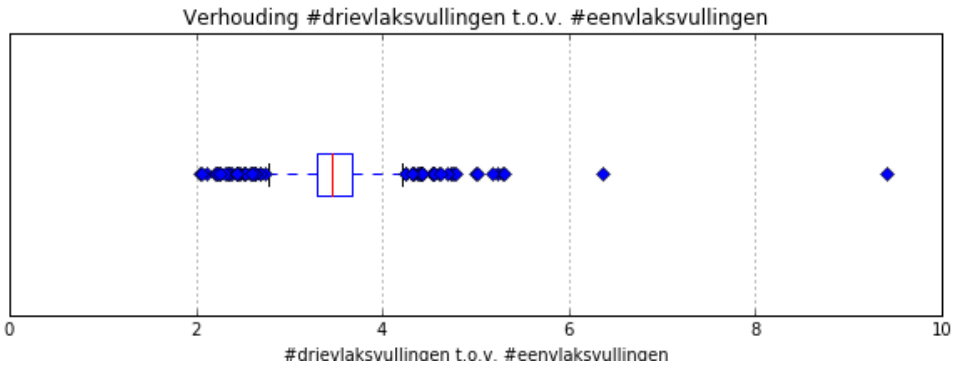
Methode T4b: Standaard Boxplot

Het ligt mogelijk niet voor de hand om een Boxplot te gebruiken voor twee afhankelijke variabelen. In het geval dat beide variabelen lineair afhankelijk zijn, kan men beide variabelen transformeren naar 1 dimensie, zoals hieronder staat afgebeeld met de test dataset van T4.



Figuur 5.22: Transformatie test dataset voor generalisatie T4

De bijhorende boxplot wordt dan:



Figuur 5.23: Resultaat boxplot voor generalisatie T4

Convolutiematrix:

		Voorspeld	
		Fraude	**Legitiem**
Werkelijk	**Fraude**	**36**	**14**
	Legitiem	**34**	**916**

Tabel 5.21: Convolutiematrix standaard Boxplot voor generalisatie T4

Methode T4c: Geoptimaliseerde Boxplot

Wanneer de parameters optimaliseren naar 1,3 × IKA (interkwartielafstand) wordt het resultaat iets beter:

		Voorspeld	
		Fraude	**Legitiem**
Werkelijk	**Fraude**	**39**	**11**
	Legitiem	**37**	**913**

Tabel 5.22: Convolutiematrix geoptimaliseerde Boxplot voor generalisatie T4

Vergelijking

Methode	Recall	Precision	F1-score
RANSAC	0.80	0.98	**0.88**
Geoptimaliseerde Boxplot	0.78	0.51	**0.62**
Standaard Boxplot	0.72	0.51	**0.60**

Tabel 5.23: Resultaten generalisatie T4

5.6.5. Generalisatie T5 - Eén cluster met twee onafhankelijke kenmerken

Generalisatie gaat T5 gaat uit van de volgende kenmerken:

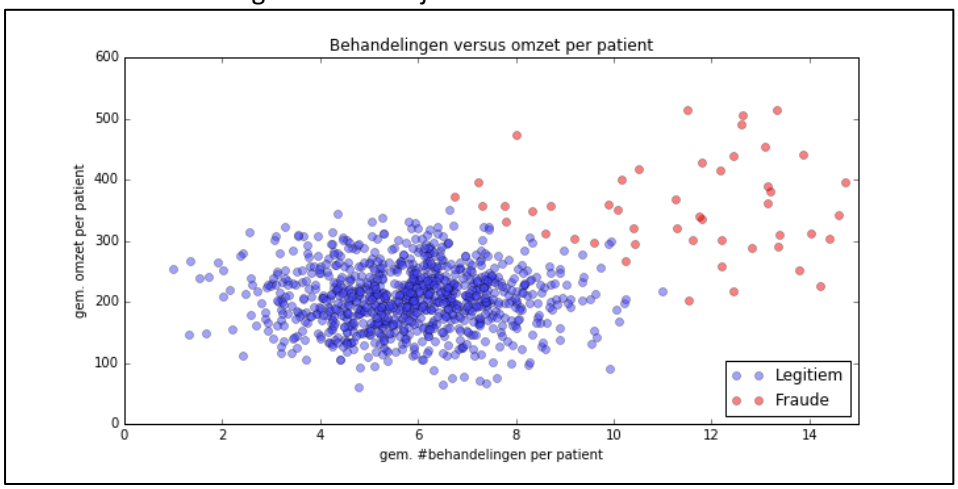
Kenmerken	Waarde
Probleem dimensie	>2
Aantal clusters	1
Distributie	Normaal verdeeld
Correlatie tussen de kenmerken	Onafhankelijk

Tabel 5.24: Kenmerken generalisatie T5

Een voorbeeld fraudescenario is GG07, maar de toepassing is erg generiek. Uitgangspunt is hierbij dat zowel het gemiddeld aantal behandelingen per patiënt, als de gemiddelde omzet per patiënt, per zorgaanbieder kunnen verschillen. Het kan echter opvallend zijn, wanneer beide kenmerken afwijken.

Test dataset

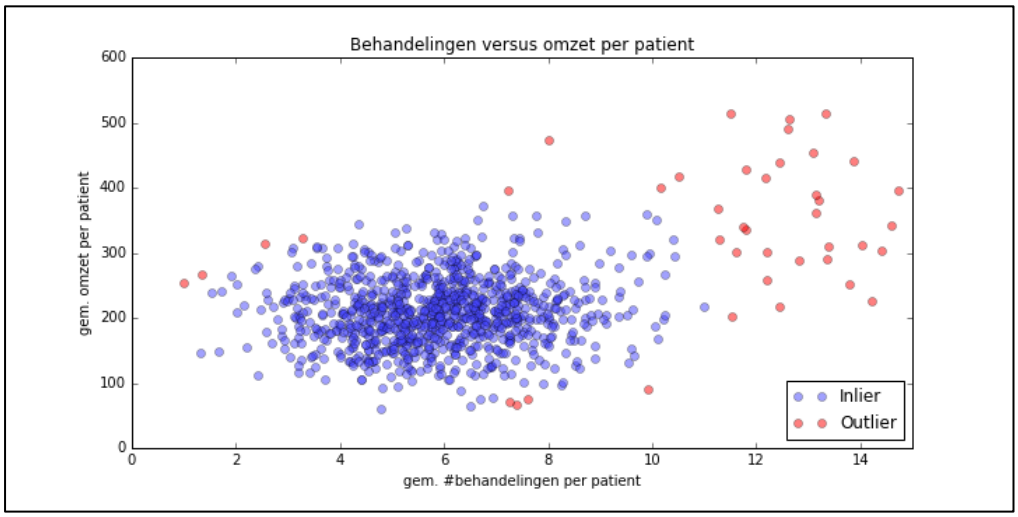
In de test dataset, zijn een deel van de fraudegevallen zo opgezet, dat deze op basis van een individueel kenmerk geen outlier zijn.



Figuur 5.24: Test dataset voor generalisatie T5

Methode T5a: Gaussian mixture model (GMM)

Het gaussian mixture model, maakt een benadering van de tweedimensionale verdeling op basis van een ellips, waarvan het middelpunt bepaald wordt op basis van de gemiddelden van beide kenmerken en de lengte, breedte en de hoek, bepaald worden door de verdeling van de data. Het gaussian mixture model is uitsluitend toepasbaar voor clusters die een elliptische vorm hebben.



Figuur 5.25: Resultaat GMM voor generalisatie T5

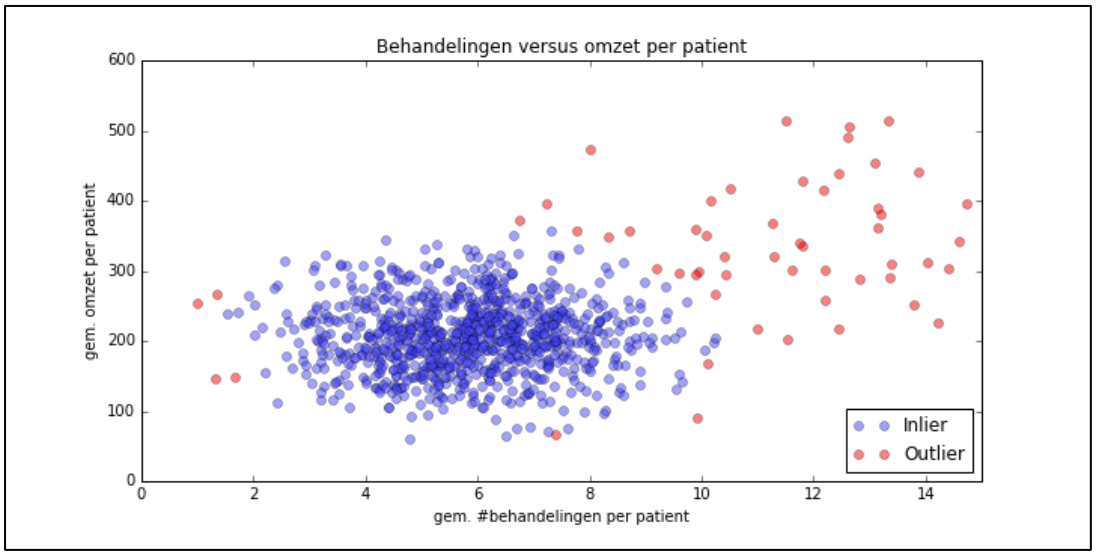
Convolutiematrix:

		Voorspeld	
		Fraude	**Legitiem**
Werkelijk	**Fraude**	**39**	**10**
	Legitiem	**10**	**941**

Tabel 5.25: Convolutiematrix GMM voor generalisatie T5

Methode T5b: DBSCAN

Onderstaande figuur geeft aan welke observaties als outlier worden gelabeld door de methode DBSCAN.



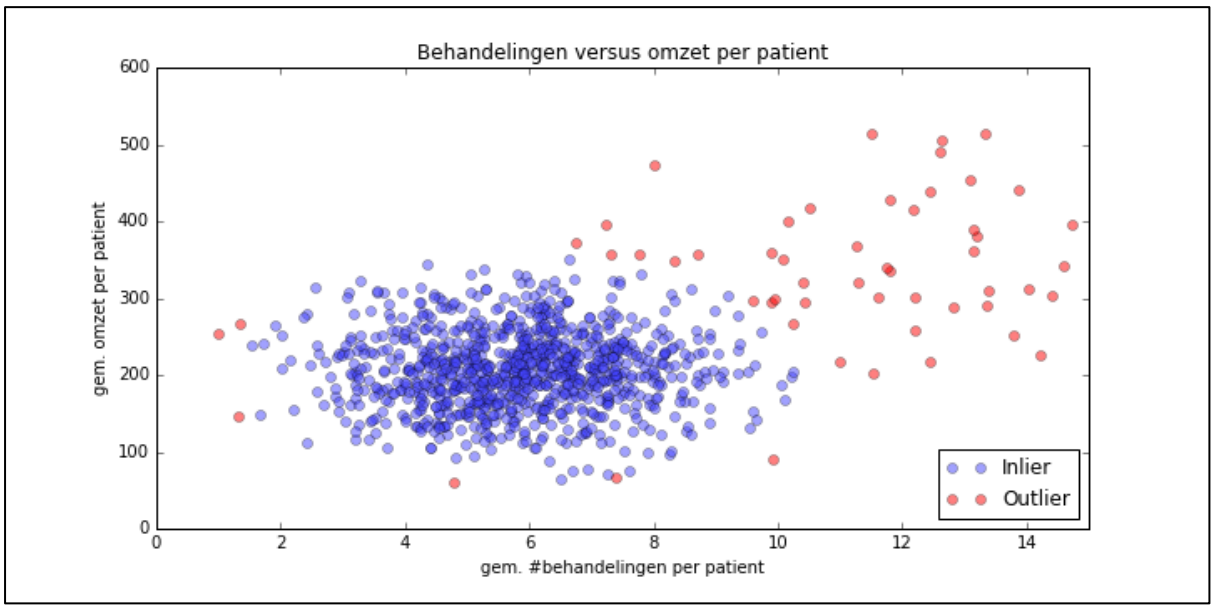
Figuur 5.26: Resultaat DBSCAN voor generalisatie T5

		Voorspeld	
		Fraude	**Legitiem**
Werkelijk	**Fraude**	**47**	**3**
	Legitiem	**10**	**940**

Tabel 5.26: Convolutiematrix DBSCAN voor generalisatie T5

Methode T5c: DeBaCl

Onderstaande figuur geeft aan welke observaties als outlier worden gelabeld door de methode DeBaCl.



Figuur 5.27: Resultaat DeBaCl voor generalisatie T5

Convolutiematrix:

		Voorspeld	
		Fraude	**Legitiem**
Werkelijk	**Fraude**	**47**	**3**
	Legitiem	**9**	**941**

Tabel 5.27: Convolutiematrix DeBaCl voor generalisatie T5

Vergelijking

Methode	Recall	Precision	F1-score
DeBaCl	0.94	0.84	**0.89**
DBSCAN	0.94	0.82	**0.88**
GMM	0.80	0.80	**0.80**

Tabel 5.28: Resultaten generalisatie T5

5.6.6. Generalisatie T6 - Meerdere clusters met twee onafhankelijke kenmerken

Generalisatie gaat T6 gaat uit van de volgende kenmerken:

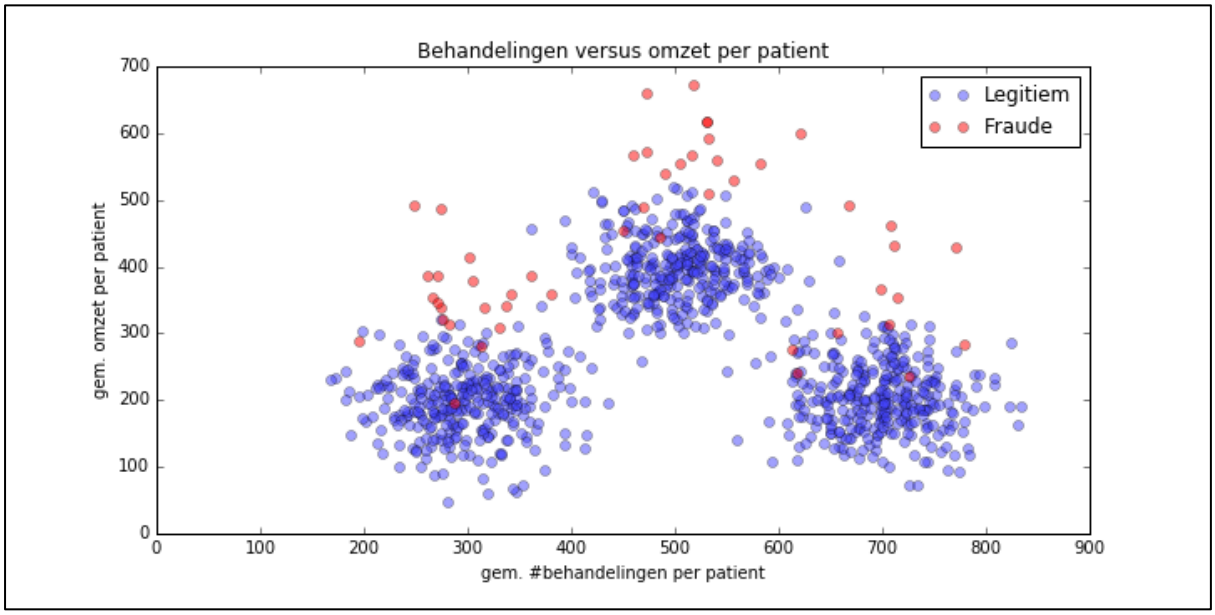
Kenmerken	Waarde
Probleem dimensie	>2
Aantal clusters	>1
Distributie	Normaal verdeeld
Correlatie tussen de kenmerken	Per cluster afhankelijk

Tabel 5.29: Kenmerken generalisatie T5

Test dataset

In de test dataset, zijn een deel van de fraudegevallen zo opgezet, dat deze op basis van een individueel kenmerk geen outlier zijn.

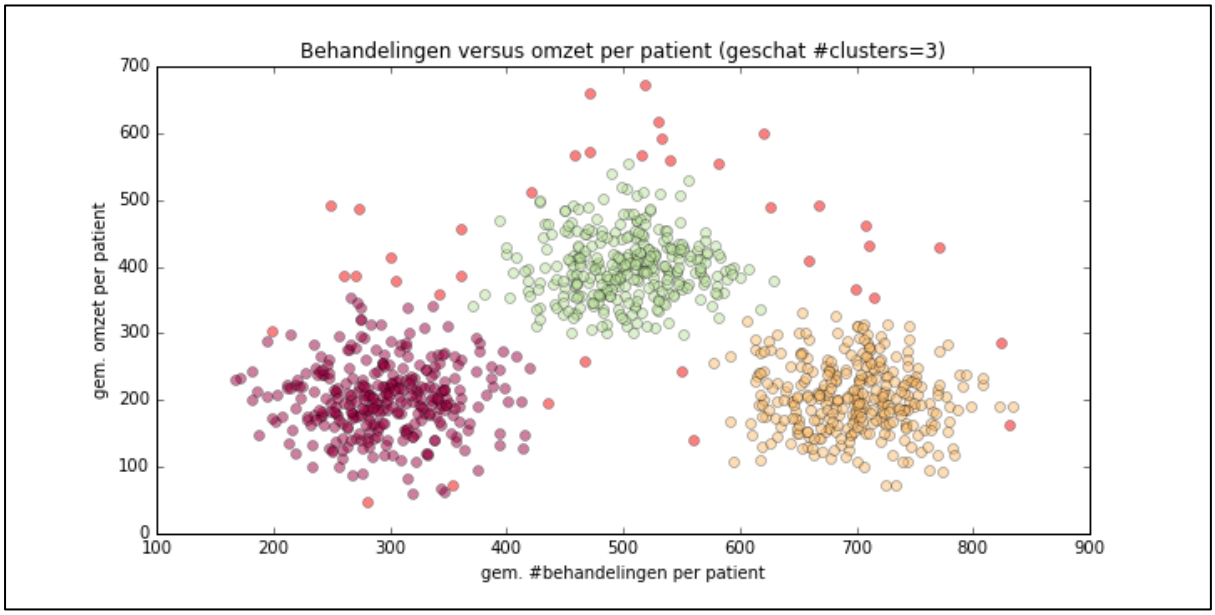
Deze fictieve testset, zou drie typen zorgaanbieders kunnen presenteren, die ieder een cluster vormen. Zoals onderstaande figuur weergeeft, zijn de fraudegevallen in dit voorbeeld lastiger te onderscheiden.



Figuur 5.28: Test dataset voor generalisatie T6

Methode T6a: DBSCAN

Onderstaande figuur geeft aan welke observaties als outlier worden gelabeld door de methode DBSCAN.



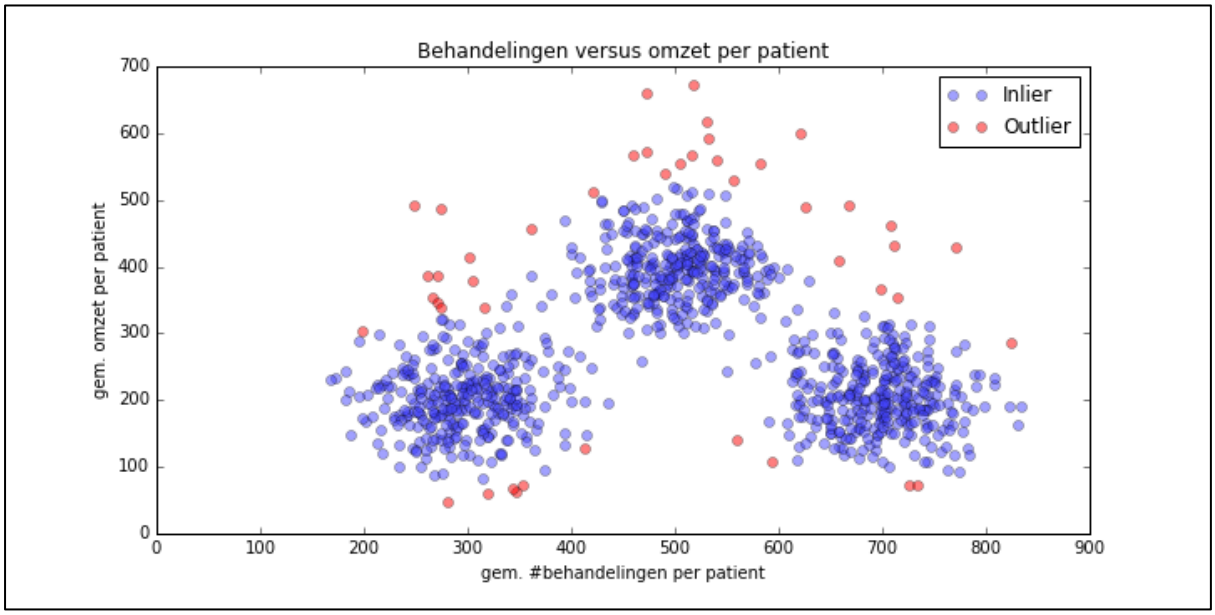
Figuur 5.29: Resultaat DBSCAN voor generalisatie T6

		Voorspeld	
		Fraude	**Legitiem**
Werkelijk	**Fraude**	**29**	**20**
	Legitiem	**22**	**929**

Tabel 5.30: Convolutiematrix DBSCAN voor generalisatie T6

Methode T6b: DeBaCl

Onderstaande figuur geeft aan welke observaties als outlier worden gelabeld door de methode DeBaCl.



Figuur 5.30: Resultaat DeBaCl voor generalisatie T6

Convolutiematrix:

		Voorspeld	
		Fraude	**Legitiem**
Werkelijk	**Fraude**	**29**	**20**
	Legitiem	**16**	**935**

Tabel 5.31: Convolutiematrix DeBaCl voor generalisatie T6

Vergelijking

Methode	Recall	Precision	F1-score
DeBaCl	0.59	0.64	**0.62**
DBSCAN	0.59	0.57	**0.58**

Tabel 5.32: Resultaten generalisatie T6

6. Resultaten

Uit het empirisch onderzoek, zijn de volgende resultaten naar voren gekomen:

Generalisatie	Recall	Precision	F1-score	Performance
T1: Normale verdeling met één kenmerk				
Geoptimaliseerde Boxplot	0.75	0.90	**0.82**	O(n)
GMM	0.75	0.88	**0.81**	O(n log n)
Standaard Boxplot	0.73	0.90	**0.80**	O(n)
T2: Normale verdeling met één kenmerk				
Geoptimaliseerde Boxplot	0.92	0.63	**0.75**	O(n)
GMM	0.69	0.62	**0.65**	O(n log n)
Standaard Boxplot	1.00	0.48	**0.64**	O(n)
T3: Gecombineerde normale verdeling met één kenmerk				
GMM	0.85	0.80	**0.83**	O(n log n)
DeBaCl	0.83	0.80	**0.82**	O(n log n)
DBSCAN	0.81	0.71	**0.76**	O(n log n)
T4: Meerdere afhankelijke kenmerken (of categorieën)				
RANSAC	0.85	0.80	**0.83**	O(n log n)
Geoptimaliseerde Boxplot	0.83	0.80	**0.82**	O(n log n)
Standaard Boxplot	0.81	0.71	**0.76**	O(n)
T5: Eén cluster met twee onafhankelijke kenmerken				
DeBaCl	0.94	0.84	**0.89**	O(n log n)
DBSCAN	0.94	0.82	**0.88**	O(n log n)
GMM	0.80	0.80	**0.80**	O(n log n)
T6: Meerdere clusters met twee onafhankelijke kenmerken				
DeBaCl	0.59	0.64	**0.62**	O(n log n)
DBSCAN	0.59	0.57	**0.58**	O(n log n)

Tabel 6.1: Kwantitatieve resultaten

Uit Tabel 6.1: Kwantitatieve resultaten, kan de onderstaande toepasbaarheidsmatrix afgeleid worden:

Generalisatie		RANSAC	GMM	Boxplot	DBSCAN	DeBaCl
T1	Normale verdeling met één kenmerk		++	++		
T2	Lognormale verdeling met één kenmerk		++	++		
T3	Gecombineerde normale verdeling met één kenmerk		++		+	++
T4	Meerdere afhankelijke kenmerken (of categorieën)	++		+		
T5	Eén cluster met twee onafhankelijke kenmerken		+/-		++	++
T6	Meerdere clusters met twee onafhankelijke kenmerken		+/-		++	++

Tabel 6.2: Toepasbaarheidsmatrix per generalisatie

Hieronder is de toepasbaarheidsmatrix in tabelvorm weergegeven

Generalisatie		Optimaal	Alternatief
T1	Normale verdeling met één kenmerk	Boxplot	GMM
T2	Lognormale verdeling met één kenmerk	Boxplot	GMM
T3	Gecombineerde normale verdeling met één kenmerk	GMM	DeBaCl OPTICS DBSCAN
T4	Meerdere afhankelijke kenmerken (of categorieën)	RANSAC	Boxplot
T5	Eén cluster met twee onafhankelijke kenmerken	DeBaCl OPTICS DBSCAN	GMM*
T6	Meerdere clusters met twee onafhankelijke kenmerken	DeBaCl OPTICS DBSCAN	GMM*

Tabel 6.3: Toepasbaarheid per generalisatie

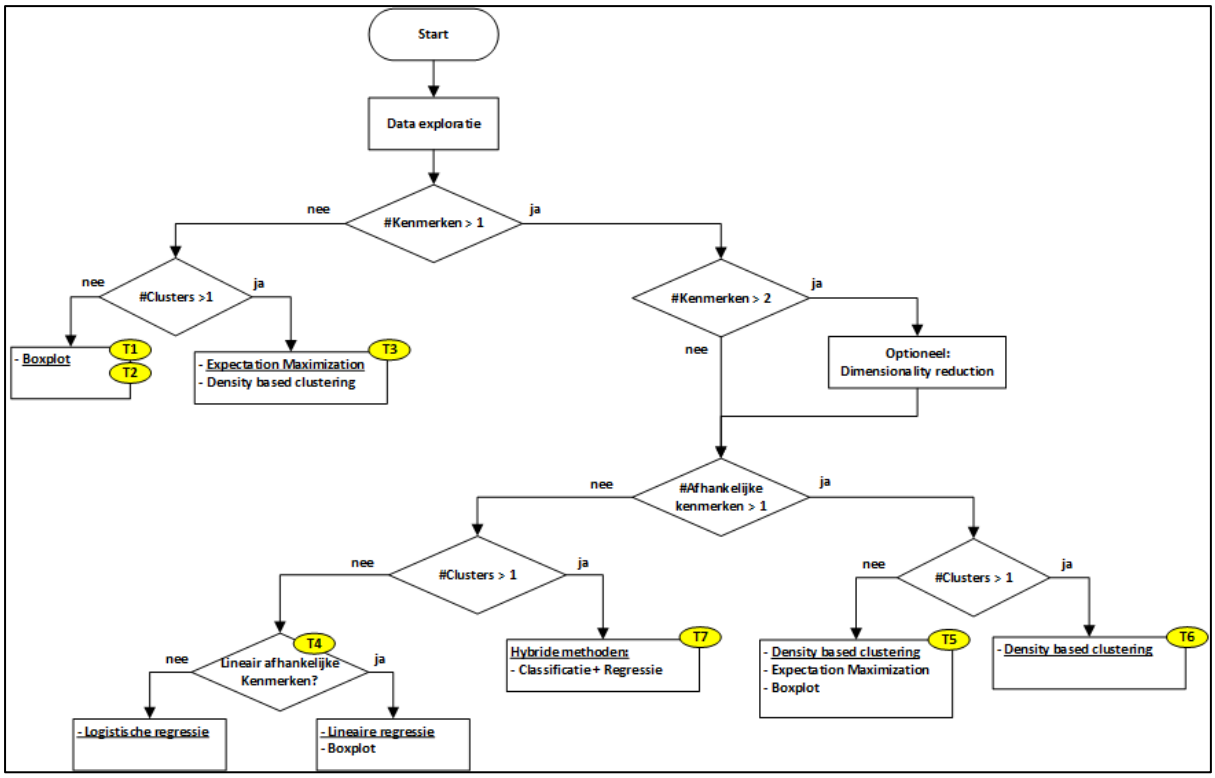
*) GMM is toepasbaar bij clusters waarvan de kenmerken normaal verdeeld zijn. Dit zijn veel voorkomende ellipsvormige clusters.

Voor zowel T1(normale verdeling) als voor T2 (lognormale verdeling), prefereert Boxplot boven GMM. Bij de toepassing van GMM in T2, dienen er meerdere gaussians toegepast te worden, om de lognormale verdeling te benaderen.

6.1.1. Keuze-framework

Door de kenmerken van de generalisaties, zoals uitgewerkt in hoofdstuk 5.3.2, als keuzes weer te geven, kan een flowchart afgeleid worden. Het keuze-framework geeft een aanbeveling over de meest geschikte outlier-detectiemethoden, waarbij de generalisaties niet expliciet meer vermeld hoeven te worden.

In onderstaande figuur, is het keuze-framework als flowchart weergegeven.



Figuur 6.1:Keuze-framework

De onderstreepte methoden zijn de voorkeurs methoden, omdat deze in het empirisch het beste resultaat geven.

Om het keuze-framework effectief toe te kunnen passen, is het belangrijk dat men de data goed kent. Het verkennen van de data, gebeurt in de stap Data Exploratie. Hierbij worden door bijvoorbeeld een data-analyst en een domein expert, de kenmerken van de data bepaald, die voor het keuze-framework van belang zijn.

Voor de bepaling van de meest geschikte methoden, zijn de volgende kenmerken van belang:

- ❑ #kenmerken: het aantal kenmerken;
- ❑ #clusters: het aantal clusters (of gaussians);
- ❑ Of de kenmerken afhankelijk zijn van elkaar;

Indien er meer dan twee kenmerken (dimensies) zijn, kan het praktisch zijn om het aantal dimensies te reduceren (dimensionality reduction) naar bijvoorbeeld twee dimensies. Een veel gebruikte methode hiervoor is Principal Component Analysis (PCA). Hierdoor is de data beter visualiseerbaar.

6.1.2. Varianten

De meeste outlier-detectiemethoden, hebben verschillende varianten met specifieke eigenschappen die de toepasbaarheid kunnen verhogen. Binnen het onderzoek, zijn wel varianten gebruikt om de effectiviteit van de methoden te bepalen, maar zijn niet alle varianten getest.

De varianten verschillen vaak erg in parameters, waarmee men de methoden kan optimaliseren.

Het is echter raadzaam om op basis van de specifieke eigenschappen van deze varianten, één of meerdere varianten te kiezen en te vergelijken.

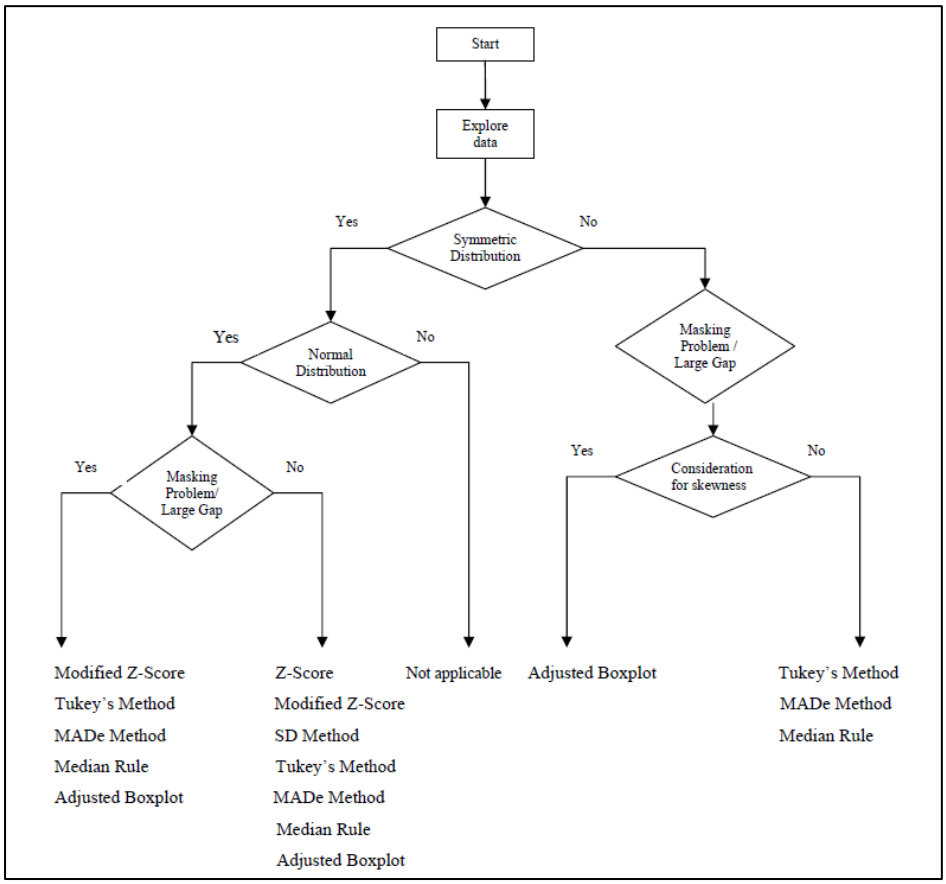
Methoden	Varianten
Expectation Maximization	GMM
Density Bases Clustering	DeBaCl OPTICS DBSCAN
Boxplot (onder T1)	Tukey's Method Median Rule MADe Method Adjusted Boxplot SD Method Z-Score Adjusted Z-Score
Regressie	RANSAC Lineaire Regressie Logistische regressie

Tabel 6.4: Varianten outlier-detectiemethoden

OPTICS

De methode OPTICS leent zich uitstekend voor interactieve analyse en bij clusters van variërende dichtheid.

Voor de verschillende Boxplot varianten bij generalisatie T1 en T2, kan gebruik gemaakt worden van de beslisboom uit het onderzoek van Songwon Seo (Songwon, 2006):



Figuur 6.2: Beslisboom voor scenario's met 1 kenmerk

7. Discussie

Hoe representatief is de testdata?

Om de effectiviteit van de outlier-detectiemethoden te beoordelen, is gebruik gemaakt van gegenereerde testdata. Hoewel er met drie zorgverzekeraars, in de praktijk binnen veel projecten samengewerkt wordt, stonden zij er niet voor open om productiedata te gebruiken voor het onderzoek. Om de verschillende methoden toch te kunnen vergelijken, is gebruik gemaakt van testdata sets. Hierbij is wel rekening gehouden met een representatieve verdeling van de data, op basis van input van domeinexperts. In de testset zijn onder ander de volgende typen outliers buiten beschouwing gelaten:

- Onopzettelijke structurele foutieve declaraties door een zorgaanbieder. Hoewel dit geen fraude is, zijn deze gevallen voor een zorgverzekeraar wel belangrijk om te bespreken;
- Afwijkingen door beperkt aantal declaraties bij een zorgaanbieder;
- Afwijkende declaratiepatronen, die verklaard kunnen worden door een (legitieme) specifieke manier van werken door een zorgaanbieder, bijvoorbeeld een bepaald specialisme of een arts die alleen op maandag werkt.

Hoewel testdata slechts een model is van de werkelijkheid, heeft het gebruik van testdata voor de vergelijking ook een aantal voordelen:

- In de testdataset, kunnen we fraudegevallen injecteren en precies analyseren welke gevallen de outlier-detectiemethode kan vinden;
- Omdat er in de test dataset minder ruis zit, worden de verschillen tussen de outlier-detectiemethode uitvergroot en beter vergelijkbaar.

Voor het bepalen van de werkelijke effectiviteit van outlier-detectiemethode, is testdata onvoldoende representatief. Voor de vergelijking van de outlier-detectiemethodes en het opzetten ven een keuze-framework, biedt het juist wel voordelen.

Tuning van parameters

Outlier-detectiemethodes hebben parameters om de werking te tunen. Hierdoor heeft men invloed op de grens waarop een observaties als outlier geclassificeerd wordt. Indien men de grens zodanig instelt, dat de kans op outliers hoger wordt, neemt vaak ook de kans op false positives toe. De outliers worden dus minder betrouwbaar. Voor een hogere betrouwbaarheid, zal men ook genoegen moeten nemen met minder outliers. Bij de vergelijking van de methoden, is de precisie(*precision*) vergeleken, bij een vast aantal outliers (*recall*).

Tijdens het onderzoek zijn verschillende parameterinstellingen uitgeprobeerd, maar het is niet uitgesloten dat de methoden door betere instellingen nog beter kunnen presteren.

Fout, regulier afwijkend gedrag of fraude

Structurele fraude manifesteert zich veelal tot afwijkend gedrag, wat weer kan leiden tot detecteerbare outliers. Het is echter niet zo dat alle outliers ook daadwerkelijk op fraude duiden.

Andere oorzaken van outliers bij zorgdeclaraties, zijn bijvoorbeeld:

- Fouten, bijvoorbeeld bij de invoer, verwerking, registratie of interpretatie;
- Afwijkende kenmerken door een beperkt aantal declaraties;
- Specialisatie van een zorgaanbieder;
- Legitiem afwijkende werkwijze van een zorgaanbieder.
- Optimalisatie van DBC-codes om zoveel mogelijk uitbetaald te krijgen. Hoewel dit grenst aan fraude, valt optimalisatie binnen de wettelijke bepalingen.

Een bijkomend probleem is dat het vaak lastig is, om fraude daadwerkelijk aan te tonen.

Clusters van fraude

Unsupervised outlier-detectiemethoden, gaan ervan uit de er onderscheid gemaakt kan worden in normaal – en afwijkend gedrag. Normaal gedrag wordt hierbij herkend doordat een cluster van observaties dicht bij elkaar liggen, of dicht bij een bepaalde (regressie-)lijn. Een observatie wordt als outlier geclassificeerd, wanneer deze te ver van een cluster of regressielijn ligt.

Het is echter niet uitgesloten dat er zorgaanbieders zijn, die op een consistente wijze frauderen, waardoor deze groep een eigen cluster vormt. Dus als een grote groep zorgaanbieders op dezelfde wijze fraude pleegt, bestaat er de kans dat dit als normaal gezien wordt.

Decentraal per zorgverzekeraar of centraal door Vektis?

De kracht van outlier-detectie, hangt erg af van het aantal observaties (zorgaanbieders) en het aantal declaraties per zorgaanbieder. Een zorgverzekeraar heeft uitsluitend toegang tot haar eigen declaraties. De declaraties van een zorgaanbieder zijn verspreid over meerdere zorgverzekeraars, waardoor een zorgverzekeraar geen totaalbeeld heeft van alle declaraties. Zeker wanneer een zorgaanbieder, samen met een vaste groep van patiënten fraude pleegt en deze patiënten jaarlijks wisselen van zorgverzekeraar.

Landelijk worden alle declaraties geregistreerd in een datawarehouse bij de organisatie Vektis. Opsporing van frauduleuze Zorgaanbieders o.b.v. afwijkende declaratiepatronen, zou het meest effectief zijn op het datawarehouse bij Vektis.

8. Conclusies en aanbevelingen

8.1. Conclusies

De centrale onderzoeksvraag is:

Welke outlier-detectiemethoden kunnen toegepast worden op welk klassen declaratiefraude scenario's?

Het afgeleide keuze-framework in 6.1.1, geeft op basis van een beperkt aantal kenmerken van het declaratiefraude scenario een duidelijke richting voor de toepasbare outlier-detectiemethoden en eventuele alternatieven.

Bij de afleiding van de kenmerken, is een aantal opvallende aspecten naar voren gekomen:

Inzichtelijk maken van de dataverdeling

Voor de keuze van de optimale outlier-detectiemethoden, is het van belang inzicht te hebben in de gegevensdistributie van de eigenschappen. Hierbij is een grafische visualisatie met behulp van een histogram of scatterplot veelal noodzakelijk.

Varianten

De meeste outlier-detectiemethode, kennen diverse varianten en implementaties. Voor een optimaal resultaat, kan men experimenteren met meerdere varianten van een bepaalde methode.

Parameter tuning bij het toepassen van een methode

De meeste methodes kennen verschillende parameters, waarmee de methode kunt tunen. Hierbij kan vaak ook een optimale verhouding tussen de *recall* en de *precision* bepaald worden.

8.2. Aanbevelingen voor vervolgonderzoek

Het onderzoek heeft als doel een keuze-framework op te stellen om zorgaanbieders op te sporen die frauderen met declaraties. Voor het keuze-framework is het om de vergelijking van de verschillende methodes, belangrijker dan de specifieke toepasbaarheid van de methodes. Het onderzoek geeft geen indicatie wat de scoringspercentages in de praktijk zullen zijn. De aanbevelingen die hieronder volgen, hebben zowel betrekking op het verbeteren van het keuze-framework, als op het uitvoeren van een praktijkonderzoek.

Voor het vervolgonderzoek kunnen zijn de test datasets in CSV formaat beschikbaar gesteld in bijlage 3.

In hoofdstuk 7 wordt een aantal beperking beschreven, die in een vervolgonderzoek ook geadresseerd kunnen worden.

Verdeling van kenmerken beoordelen

De kenmerken van de zorgaanbieders zijn veelal afgeleide gemiddelden, zoals:

- ❑ Aantal behandelingen per patiënt;
- ❑ Omzet per patiënt;
- ❑ Verhouding #drievlaksvullingen t.o.v. #eenvlaksvullingen.

Outlier-detectie kan mogelijk verbeterd worden, door niet alleen uit te gaan van de gemiddelde waarde van een kenmerk, maar ook rekening te houden met de spreiding van de data. Uitgangspunt is dat het bijvoorbeeld uitmaakt hoe de omzet per patiënt bij een zorgaanbieder met een gemiddelde omzet van €200,- per patiënt verdeeld is. Mogelijk is bij 95% van de patiënten de omzet lager dan €50,- en bij een aantal hoger den €1.000,-. Wellicht is het raadzaam om deze extremen zelfs uit te sluiten als observatie omdat dit een verstoord beeld geeft.

Afwijkend gedrag in de tijd

Het kan interessant zijn om de hypothese te onderzoeken dat zorgaanbieders die declaratiefraude plegen, afwijkend gedrag in de tijd vertonen. Het idee hierachter is dat er in de tijd fluctuaties zullen zijn omdat het lastig is om frauduleuze declaraties consistent in de tijd te verdelen. Daarnaast is het aannemelijk dat meest buitensporige zorgaanbieders, steeds meer zullen gaan frauderen, wat tot een verschuiving in kengetallen kan leiden.

Combineren van kenmerken

Het onderzoek richt zich op het vinden van frauduleuze zorgaanbieders op basis van een aantal generieke kenmerken, welke afgeleid zijn uit een vooraf gedefinieerde scenario's. Er zijn echter verschillende fraudescenario's detecteerbaar op basis van dezelfde kenmerken, zelfs scenario's die tot op heden nog onbekend waren.

Kenmerken uit verschillende scenario's kunnen met bijvoorbeeld neutrale netwerken gecombineerd worden, om afwijkend gedrag te detecteren.

Combineren van kansen

Er zijn verschillende modellen waar verdachte zorgaanbieders uit naar voren kunnen komen. Het is aannemelijk dat de kans op afwijkend gedrag groter is, bij zorgaanbieders die bij meerdere modellen als outlier naar voren komt. Bijvoorbeeld scenario's EF04 en EF08.

Productie data

De toepasbaarheid van outlier-detectiemethoden om declaratiefraude op te sporen, zou in de praktijk getoetst kunnen worden op een productie dataset van een (grote) zorgverzekeraar of landelijk bij Vektis. Dit zou gecombineerd kunnen worden met de analyse geconstateerde outliers door fraude-specialisten.

Optimalisatie o.b.v. resultaten

De methoden zijn nu volledig unsupervised. Er kan met finetuning van de methode-parameters, mogelijk rekening gehouden worden met resultaten (zoals *false positives*) uit opvolgend fraudeonderzoek, maar de voorgestelde methodes, leren niet van resultaten. Hier is aanvullend onderzoek voor nodig.

Mate van outlierness

In het onderzoek, wordt een observatie geclassificeerd als wel of geen outlier. De meeste methoden ondersteunen ook de mogelijkheid om de zogenaamde *outlierness* te bepalen. Bij clustering is de *outlierness* hoger, naarmate de outlier verder van een cluster verwijderd is. De praktische toepasbaarheid van outlier detectie-methodes, zal stijgen wanneer de kans op fraude(*outlierness*) en de impact(potentieel verhaalbare schade) inzichtelijk gemaakt kunnen woorden.

9. Reflectie

9.1. Kwaliteit van het onderzoek

Het onderzoek heeft mijns inziens een bruikbaar framework opgeleverd, waarmee een belangrijke stap gezet is om declaratiefraude in de zorg tegen te gaan. Daarmee is het doel van het onderzoek gerealiseerd. In het onderzoek, zijn niet alle varianten van methoden vergeleken en getoetst. Het keuze-framework, geeft via de geadviseerde methoden aan welke varianten mogelijk interessant zijn voor bepaalde scenario's.

De outlier-detectiemethoden zijn voor de scenario's geselecteerd op basis van hun eigenschappen volgens de literatuur. Daarnaast zijn deze methoden getoetst met testdata, waaruit gebleken is dat de methoden doen wat ze theoretisch zouden moeten doen.

Een belangrijke vraag die echter nog openstaat, is in hoeverre outlier-detectiemethoden daadwerkelijk toepasbaar zijn binnen de zorg.

Om de methoden te toetsen, wordt gebruik gemaakt van test datasets, waarbij per generalisatie een test dataset opgezet wordt van een representatief fraudescenario. Hierdoor worden de grafieken concreter en beter te interpreteren. De dataset is gegenereerd op basis van input van domein experts. Omdat niet getoetst kan worden hoe representatief deze gegenereerde data is, kan uit het onderzoek de effectiviteit van een methode ook niet met een redelijke betrouwbaarheid vastgesteld worden. Het onderzoek richt zich echter op het vergelijken van de toepasbaarheid, waarvoor de test dataset wel voldoende representatief moet zijn. Het gaat er hierbij dus niet om dat de test dataset, voldoende representatief is voor de werkelijke productiedata, maar voldoende representatief voor de gehele generalisatie.

9.2. Proces

Zorgfraude is een onderwerp dat zich goed leent voor onderzoek. Het opsporen van zorgfraude staat in Nederland namelijk nog in de kinderschoenen, terwijl in het bredere verzekeringsdomein al veel onderzoeken te vinden zijn over claim fraude. Het declaratieproces is ook redelijk eenvoudig en goed uit te leggen aan leken. Iedereen is immers betrokken en gebaat bij goede zorg.

Het initiële doel van de afstudeeropdracht was om aan te tonen dat machine learning toepasbaar is voor het opsporen van declaratiefraude in de Zorg. Het idee was om samen met een zorgverzekeraar, machine learning in te zetten voor enkele praktijkscenario's op werkelijke productiedata. Ik heb meerdere zorgverzekeraars gevraagd om hieraan mee te werken, maar zij konden geen productiedata ter beschikking stellen en waren ook erg terughoudend met het vrijgeven van informatie omtrent zorgfraude.

Dit was een enorme tegenvaller. Het onderzoek voortzetten met test datasets, was onvoldoende representatief. Daarop is gekozen om de opdracht aan te passen in het opstellen van een keuze-framework voor outlier-detectiemethoden. Test dataset zijn namelijk zeer geschikt om verschillende methoden te vergelijken. Deze nieuwe doelstelling legde meer focus op de opdracht, waardoor de vervolgacties redelijk voorspelbaar werden.

Bij het literatuuronderzoek naar de outlier-detectiemethoden, maar vooral ook bij het empirisch onderzoek naar de verschillende varianten en implementaties, bleken er te veel varianten te zijn om allemaal te vergelijken en te testen. In het onderzoek heb ik daarom de nadruk gelegd op de methoden, met één of twee varianten, die in Python zijn geïmplementeerd. Hierdoor konden de methoden goed geëvalueerd worden in de tool Jupyter Notebook.

Het onderzoek heb ik uitgevoerd naast mijn gezin en mijn fulltimebaan als projectmanager. Bij aanvang van het onderzoek, leek het erop dat er een rustige periode aan zou komen waarin er voldoende tijd zou zijn voor het onderzoek. Dit liep helaas echter niet zo omdat er direct al reorganisatie aan kwam en een aantal grote projecten opgestart werd, waar ik verantwoordelijk

voor was. Hierdoor bleken verschillende planningen niet realistisch en zijn ook niet gehaald. Pas eind 2016 was er weer tijd om het onderzoek goed op te pakken en af te ronden.

Het advies van mijn begeleider om wekelijks minimaal één pagina te schrijven en zo toch voortgang te houden, heeft daarbij enorm geholpen.

Referenties

Abraham, B., & Box, G. E. (1979). Bayesian analysis of some outlier problems in time series. *Biometrika 66*, 229-236.

Abraham, B., & Chuang, A. (1989). Outlier detection and time series modeling. *Technometrics 31*, 229-236.

Agrawal, S., & Agrawal, J. (2015). Survey on Anomaly Detection using Data Mining Techniques. *Procedia Computer Science 60*, 708 – 713.

Aitkin, M., & Wilson, G. (1980). Mixture models, outliers, and the EM algorithm. *Technometrics, 22*, 325-331.

Anderson, D., Frivold, T., Tamaru, A., & Valdes, A. (1994). *Next-generation intrusion detection expert system (nides), software users manual, beta-update release. Tech. Rep. SRI-CSL-95-07.* Computer Science Laboratory, SRI International.

Ankerst, M., Breunig, M., Kriegel, H., & Sander, J. (1999). OPTICS: Ordering Points To Identify Clustering Structure,. *IEEE Journal on Selected Areas in Communications, Special Issue on Intelligent Signal Processing for Communications 12*.

Anscombe, F. J., & Guttman, I. (1960). Rejection of outliers.

Chandola, V., Banerjee, A., & Kumar, V. (2007). *Outlier Detection: A Survey.*

Chandola, V., Banerjee, B., & Kumar, V. (2009). Anomaly Detection : A Survey. *ACM Computing Surveys*, 1-72.

Dasgupta, D., & Nino, F. (2000). A comparison of negative and positive selection algorithms in novel pattern detection. *In Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics. Vol. 1.* Nashville.

Eskin, E. (2000). Anomaly detection over noisy data using learned probability distributions. *In Proceedings of the Seventeenth International Conference on Machine Learning* (pp. 255-262). Morgan Kaufmann Publishers Inc.

Eskin, W. L., & Stolfo, S. (2001). Modeling system call for intrusion detection using dynamic window sizes. *In Proceedings of DARPA Information Survivability Conference and Exposition.*

Ester, M., Kriegel, H., Sander, J., & Xu, X. (1996). A density-based algorithm for discoveringclusters in large spatial databases with noise. *Proceedings of Second International Conference on Knowledge Discovery and Data Mining* (pp. 226–231). Portland, Oregon: Eds. AAAI Press.

Fawcett, T., & Provost, F. (1999). Activity monitoring: noticing interesting changes in behavior. *In Proceedings of the 5th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 53-62). ACM Press.

Fischler, M., & Bolles, R. (1981). Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM, 24(6)*, 381–395.

Fox, A. J. (1972). Outliers in time series. *Journal of the Royal Statistical Society. Series B(Methodological) 34*, 350-363.

Gee, J., & Button, M. (2015). *The financial cost of healthcare fraud 2015.* PKF Littlejohn LLP.

Gibbons, R. D. (1994). *Statistical Methods for Groundwater Monitoring.* John Wiley & Sons, Inc.

Grubbs, F. (1969). Procedures for detecting outlying observations in samples. *Technometrics 11*, 1-21.

Helman, P., & Bhangoo, J. (1997). A statistically based system for prioritizing information exploration under uncertainty. *In IEEE International Conference on Systems, Man, and Cybernetics. Vol. 27*, pp. 449-466. IEEE.

Ilgun, K., Kemmerer, R., & Porras, P. (1995). State transition analysis: A rule-based intrusion detection approach. *Transactions on Software Engineering 21* (pp. 181–199). IEEE.

Jain, A. K., & Dubes, R. C. (1988). *Algorithms for Clustering Data.* Prentice-Hall, Inc.

Jain, K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters (Volume 31, Issue 8)*, 651-666.

Javitz, H. S., & Valdes, A. (1991). The sri ides statistical anomaly detector. *In Proceedings of the 1991 IEEE Symposium on Research in Security and Privacy.* IEEE Computer Society.

Kent, B., Rinaldo, A., & Verstynen, T. (2013). DeBaCl: A Python Package for Interactive DEnsity-BAsed CLustering. *Journal of Statistical Software*.

Knorr, E. M., & Ng, R. T. (1998). Algorithms for mining distance-based outliers in large datasets. *In Proceedings of the 24rd International Conference on Very Large Data Bases* (pp. 392-403). Morgan Kaufmann Publishers Inc.

Laleh, N., & Azgomi, M. A. (2009). A Taxonomy of Frauds, Fraud Detection Techniques.

Laurikkala, J., Juhola1, M., & Kentala., E. (2000). Informal identifcation of outliers in medical data. *In Fifth International Workshop on Intelligent Data Analysis in Medicine and Pharmacology*, 20-24.

Limpert, E., Stahel, W., & Abbt, M. (2001). Log-normal Distributions across the Sciences: Keys and Clues. *ioScience, Vol. 51, No. 5,*.

Mitchell, T. M. (1997). *Machine Learning.* McGraw-Hill Higher Education.

NZa. (2014). *Rapport Onderzoek Zorgfraude.* Opgehaald van https://www.nza.nl/1048076/1048181/Rapport_Onderzoek_zorgfraude__update.pdf

Otey, M. E., Ghoting, A., & Parthasarathy, S. (2006). Fast distributed outlier detection in mixed-attribute data sets. *Data Min. Knowl. Discov. 12*, 2-3, 203-228.

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, 45-78.

Rosner, B. (1983). Percentage points for a generalized esd many-outlier procedure. *Technometrics 25*, 165-172.

Salvador, S., & Chan, P. (2003). Learning states and rules for time-series anomaly detection. *Tech. Rep. CS-2003-05, Department of Computer Science, Florida Institute of Technology Melbourne FL 32901*.

Smyth, P. (1994). Markov monitoring with unknown states. *IEEE Journal on Selected Areas in Communications, Special Issue on Intelligent Signal Processing for Communications 12*, 1600–1612.

Songwon, S. (2006). *A Review and Comparison of Methods for Detecting Outliers in Univariate Data Sets.* University of Pittsburgh.

Sparrow, M. (2000). *License to Steal: How Fraud Bleeds America's Health Care System.* Westview press.

Stefansky, W. (1972). Rejecting outliers in factorial designs. *Technometrics 14*, 469-479.

Vapnik, V. N. (1995). The nature of statistical learning theory.

Zhang, J. (2013). Advancements of Outlier Detection: A Survey. *ICST Transactions on Scalable Information Systems*.

Zorgverzekeraars Nederland. (2017). Opgehaald van Zorgverzekeraars Nederland: https://www.zn.nl/350584836/Feiten-en-cijfers

Bijlage 1: Fraude scenario's

Geestelijke Gezondheidszorg

Nr.	Scenario	Fraude categorie	Kenmerken	Type	Voorbeeld
GG01	Overbehandeling door meer zorg te leveren.	Overbehandelen	#sessies / patiënt #aangesloten sessies / patiënt #sessies / behandeling	T2	Een patiënt is al een aantal maal bij een psychiater geweest voor de behandeling van zijn depressie. De psychiater is medisch gezien uitbehandeld, maar besluit toch nog een tweetal extra sessies met deze patiënt te doen zodat de geleverde zorg leidt tot een DBC met een hoger bedrag dan wanneer hij deze sessies niet zou hebben gedaan.
GG02	Overbehandeling door een patiënt onnodig op te nemen.	Overbehandelen	#klinische behandelingen/patiënt	T3	Een patiënt komt voor een behandeling een aantal maal bij een GGZ-instelling. Binnen deze instelling is er ook ruimte om een patiënt op te nemen en klinisch te behandelen. Hoewel hiervoor de noodzaak ontbreekt, neemt de GGZ-instelling de patiënt een dag op om zo een duurder product in rekening te kunnen brengen.
GG03	Overbehandeling door een patiënt te duur op te nemen.	Overbehandelen	#patiënten/verblijfscategorie	T4	Sinds 2012 kent de productstructuur verblijf een indeling naar zwaarte. Deze indeling biedt vrijheid voor zorgaanbieders om een patiënt zelf in te delen in een verblijfscategorie. Dit heeft als risico dat patiënten in een zwaardere en dus duurdere categorie worden ingedeeld.
GG04	Onderbehandeling door minder zorg te leveren.	Onderbehandelen	#sessies / patiënt	T2	Een patiënt is al een aantal maal bij een psychiater geweest voor de behandeling van zijn depressie. De psychiater is medisch gezien nog niet uitbehandeld, maar ziet dat hij aan het begin van een tijdsklasse zit. Dat betekent dat hij na 5 extra sessies hetzelfde bedrag in rekening kan brengen als wanneer hij deze sessies niet doet. Hij besluit de sessies niet te doen en er is daarmee sprake van onderbehandeling.
GG05	Onderbehandeling door de inzet van lager gekwalificeerd personeel.	Onderbehandelen	Buiten scope: enquête / evaluatie	N.v.t.	Het bedrag dat een GGZ instelling in rekening mag brengen is gebaseerd op een verdeling in de verschillende beroepen die een patiënt behandelen. Een GGZ instelling kan er voor kiezen relatief veel beroepen in te zetten met een relatief lage kwalificatie. Ook worden ervaringsdeskundigen ingezet. In dat geval krijgt een patiënt mogelijk niet de kwalitatief juiste behandeling.
GG06	Het in rekening brengen van zorg die niet geleverd is.	Spookzorg	#behandeling / patiënt Enquête / Evaluatie	T2	Een jonge patiënt komt met zijn ouders bij de psychiater en wordt behandeld voor een depressie. Naast het in rekening brengen van deze behandeling brengt de psychiater ook voor elk van de ouders een behandeling in rekening terwijl de behandeling niet op hen gericht is.
GG07	Het in rekening brengen van zorg die deels niet geleverd is.	Spookzorg	#behandeling / patiënt #DBC / patiënt € / patiënt Enquête / Evaluatie	T2	Een patiënt komt bij de psychiater en wordt behandeld voor een depressie. De psychiater brengt een DBC in rekening met daarop ook farmacotherapie terwijl deze nooit geleverd is.
GG08	Het declareren van duurdere behandelingen	Upcoding	Indirecte tijd / contacttijd	T2	Een aanbieder werkt met normtijden. Hij rekent voor elk consult in zijn registratie standaard 1 uur directe contacttijd met de patiënt en een half uur indirecte tijd voor de voorbereiding. Zijn afspraken zijn echter altijd een uur, waardoor hij consequent meer tijd in rekening brengt dan dat hij levert.
GG09	Het declareren van duurdere behandelingen	Upcoding	€ / 1e consult tijd / 1e consult	T2	Een aanbieder ziet een patiënt voor de eerste keer. Hij registreert hiervoor een intake, een consult, het voorbereiden van het gesprek en het vragen van feedback aan een andere behandelaar. Nadat deze patiënt voor zijn eerste consult is

					geweest staat de teller al op 240 minuten, terwijl de aanbieder slechts anderhalf uur heeft besteed.
GG10	Het declareren van duurdere behandelingen	Upcoding	Enquête / Evaluatie	N.v.t.	Een aanbieder heeft groepsbehandelingen. In deze sessie van 2 uur gaat hij met 10 van zijn patiënten tegelijk in gesprek. Hij hoort dit in rekening te brengen als 12 minuten per patiënt, maar hij rekent aan elke patiënt vervolgens 2 uur toe.
GG11	Het meerdere malen in rekening brengen van dezelfde zorg	Dubbele bekostiging	€ GGZ / € AWBZ % patiënten met zowel GGZ als AWBZ	T4 T2 (T1)	Een patiënt is al in behandeling bij een GGZ instelling binnen de AWBZ. De instelling besluit dezelfde zorg echter ook in rekening te brengen via een DBC in de curatieve GGZ. Dezelfde zorg wordt nu op twee manieren bekostigd.
GG12	Het meerdere malen in rekening brengen van dezelfde zorg	Dubbele bekostiging	#DBC / patiënt #(dossiers met meerdere DBC) / patiënt	T2	Een psychiater behandelt een patiënt voor zijn depressie. Hij opent een DBC hiervoor, maar opent tegelijk ook een DBC voor een angststoornis. Er vindt 1 behandeling plaats, maar deze wordt via twee DBC's in rekening gebracht.
GG13	Het meerdere malen in rekening brengen van dezelfde zorg	Dubbele bekostiging	%patiënten met meerdere psychiaters	T2 (T1)	Een psychiater behandelt een patiënt voor zijn depressie. Hij opent een DBC hiervoor, en vraagt een collega om een consult. Deze collega zou hiervoor een vergoeding moeten krijgen van de psychiater die de DBC heeft geopend. Dit heet een intercollegiaal consult. De bevraagde collega opent echter zelf ook een DBC voor deze patiënt. Er vindt 1 behandeling plaats, maar deze wordt via twee DBC's in rekening gebracht.
GG14	Het registreren van een andere diagnose.	Onverzekerde zorg	% onverzekerde zorg (o.b.v. €) % onverzekerde zorg (o.b.v. #)	T2 (T1)	De behandeling van een aanpassingsstoornis is onverzekerde zorg. Als het om een dergelijke stoornis gaat, kan een zorgaanbieder ervoor kiezen om de behandeling onder een andere aandoening te registreren zodat de zorg wel via de zorgverzekeraar wordt vergoed.
GG15	Het splitsen van de factuur.	Opknippen	#trajecten per patiënt (binnen x maanden)	T2	Een patiënt is 10 keer bij de psychiater geweest. De psychiater dient nu 1 behandeltraject in rekening te brengen. De vergoeding voor het in rekening brengen van 2 trajecten van 5 behandelingen is echter hoger. De psychiater brengt daarom 2 behandelingen in rekening.

Extramurale Farmaceutische zorg

Nr.	Scenario	Fraude categorie	Kenmerken	Type	Voorbeeld
EF01	Het onnodig vaak voorschrijven van een geneesmiddel	Overbehandelen	€ Farmacie / patiënt	T2	Een apotheekhoudende huisarts schrijft bovengemiddeld veel geneesmiddelen voor. Zijn patiënten komen het geneesmiddel vervolgens ook weer bij hem halen.
EF02	Onnodig (veel) verstrekken van een geneesmiddel buiten weten van de apotheker om	Overbehandelen	Buiten scope: patiënt	N.v.t.	Een verslaafde patiënt krijgt op een briefje een recept mee voor pijnstillers. Hij kopieert vervolgens dit briefje en gaat bij alle apotheken in de regio langs om op deze manier veel meer dan de voorgeschreven hoeveelheid op te halen.
EF03	Minder eenheden afleveren dan nodig.	Opknippen	#bezoeken / patiënt	T2	Een apotheker levert een patiënt minder eenheden van het geneesmiddel dan voorgeschreven door een arts of overeengekomen met een zorgverzekeraar. De patiënt moet hierdoor nog eens terugkomen en de apotheker kan hierdoor nogmaals een prestatie in rekening brengen. Op het moment dat de patiënt onnodig terugkomt, is er niet langer sprake van onderbehandeling, maar van opknippen.
EF04	Meer eenheden afleveren dan nodig.	Overbehandelen	€ / patiënt eenheden / patiënt (per geneesmiddel)	T2	Een apotheker kan een patiënt ook meer eenheden van het geneesmiddel geven. In dat geval is er sprake van ondoelmatig geleverde zorg.
EF05	eenheden / patiënt (per geneesmiddel)"		#bezoeken / patiënt	T2	Een patiënt komt bij de apotheek voor zijn geneesmiddelen. De apotheek brengt het uitgeven van de geneesmiddelen en de geneesmiddelen zelf in rekening.

					Vervolgens dient hij deze factuur elke week in bij de verzekeraar terwijl de patiënt niet meer is geweest.
EF06	Het in rekening brengen van zorg die niet geleverd is.	Spookzorg	%magistrale bereidingen	T2 (T1)	Een patiënt komt bij de apotheek om zijn geneesmiddelen op te halen. De apotheker brengt een toeslag voor magistrale bereiding in rekening terwijl hij het geneesmiddel niet zelf bereid heeft.
EF07	Het in rekening brengen van zorg die deels niet geleverd is.	Spookzorg	%toeslag voor eerste uitgifte	T2 (T1)	Een patiënt komt bij de apotheek om zijn geneesmiddelen op te halen. De apotheker brengt een toeslag voor eerste uitgifte in rekening terwijl hij het geneesmiddel in de 12 maanden ervoor al meerdere malen aan deze patiënt heeft verstrekt.
EF08	Het in rekening brengen van zorg die deels niet geleverd is.	Spookzorg	#geneesmiddelen / patiënt (per geneesmiddel)	T2	Een patiënt haalt een doosje met 20 pillen bij de apotheker. De apotheker factureert echter een doosje met 40 pillen bij de verzekeraar.
EF09	Het declareren van duurdere behandelingen.	Upcoding	# bijzondere bereiding / reguliere magistrale bereiding % bijzondere bereiding	T2	Een apotheker heeft voor een patiënt een reguliere magistrale bereiding gedaan. Hij declareert deze echter als een bijzondere bereiding tegen een hoger tarief
EF10	Het declareren van duurdere behandelingen.	Upcoding	%ANZ toeslagen (meerdere klassen)	T3	Een apotheker die de ANZ (avond nacht en zondag) toeslag vanwege dienstwaarneming in rekening brengt, terwijl het geen dienstwaarneming betreft
EF11	Dubbele bekostiging	U-bochtconstructie	Buiten scope: enquête / evaluatie	N.v.t.	Een vrouw komt in het ziekenhuis voor het laten plaatsen van een spiraaltje. De behandelend arts vraagt de vrouw om eerst zelf een spiraaltje aan te schaffen via de apotheek. Vervolgens plaatst de arts dit spiraaltje. Zowel de apotheek als de arts/het ziekenhuis ontvangt nu een vergoeding voor het spiraaltje.
EF12	Het in rekening brengen van een andere prestatie om een vergoeding te krijgen.	Onverzekerde zorg	%onverzekerde geneesmiddelen (o.b.v. €) %onverzekerde geneesmiddelen (o.b.v. #)	T2 (T1)	Vanaf 2012 zijn er nieuwe prestaties. Een patiënt komt bij de apotheek voor een advies over de ziekterisico's bij het reizen. Dit is een prestatie die niet tot het verzekerde pakket behoort. De apotheker declareert een terhandstelling. Dit is verzekerde zorg waardoor de patiënt niet zelf hoeft te betalen.
EF13	Het in rekening brengen van een andere prestatie om een vergoeding te krijgen.	Onverzekerde zorg	%onverzekerde geneesmiddelen (o.b.v. €) %onverzekerde geneesmiddelen (o.b.v. #)	T2 (T1)	Een patiënt komt bij de apotheker om Viagra te halen. Viagra behoort niet tot het verzekerde pakket en de patiënt dient dit dan ook zelf af te rekenen. Hij vraagt echter om het geneesmiddel als pijnstiller in rekening te brengen zodat het wel wordt vergoed. De apotheker doet dit vervolgens.

Huisartsenzorg

Nr.	Scenario	Fraude categorie	Kenmerken	Type	Voorbeeld
HZ01	Het in rekening brengen van niet-geleverde zorg aan ingeschreven patiënten.	Spookzorg	%patiënten met consult	T2 (T1)	Een huisarts brengt jaarlijks bij al zijn ingeschreven consumenten standaard een consult in rekening zonder dat de consument daadwerkelijk bij de huisarts is geweest.
HZ02	Het vaker dan nodig leveren van een consult.	Overbehandelen	%patiënten met herhaaldelijke consults (>10) per jaar	T2 (T1)	Een huisarts heeft een aantal patiënten die elke week langs komen voor een consult. De huisarts voert deze consulten ook daadwerkelijk uit, maar er is eigenlijk geen medische noodzaak voor deze consulten. Het betreft hier ondoelmatige zorg.
HZ03	Het vaker doorverwijzen dan nodig.	Overbehandelen	%doorverwijzingen naar apotheek	T2 (T1)	Een huisarts verwijst al zijn patiënten door naar de apotheek. Zij nemen daar geneesmiddelen af. Het was echter niet nodig dat alle patiënten geneesmiddelen ontvingen. Ook hier is sprake van ondoelmatige zorg, ofwel verspilling van zorggelden.

HZ04	Het in rekening brengen van een meer uitgebreide prestatie dan geleverd	Upcoding	%langdurende consults	T2 (T1)	Een patiënt komt bij de huisarts en krijgt een regulier consult. De huisarts plant standaard patiëntenstromen van 10 minuten in, maar het consult loopt enkele minuten uit. De arts declareert een langdurend consult. De NZa heeft de grens tussen een kortdurend en langdurend consult echter op 20 minuten gesteld.
HZ05	Het in rekening brengen van een meer uitgebreide prestatie dan geleverd	Upcoding	% M&I chirurgie / per patiënt	T3	Een patiënt komt bij de huisarts voor de verwijdering van een kleine wrat. De huisarts declareert een M&I chirurgie. Dit is niet toegestaan.
HZ06	Het in rekening brengen van een andere prestatie dan geleverd	Upcoding	% M&I chirurgie / per patiënt %herhaalmedicatie	T3	Een huisarts kan de M&I-verrichting 'reizigersadvisering' declareren, terwijl de patiënt alleen vraagt om herhaalmedicatie voor malariamedicatie en geen uitgebreid advies krijgt.
HZ07	Het in rekening brengen van een andere prestatie dan geleverd	Upcoding	Buiten scope: enquête / evaluatie	N.v.t.	Het kan voorkomen dat huisartsen bij simpele vragen aan de assistent telefonische of emailconsulten rekenen, Dit is geen 'vervanger van spreekuurconsult'.
HZ08	Het in rekening brengen van een andere prestatie dan geleverd	Upcoding	%telefonisch consult	T3	Het is met ingang van 2010 niet meer toegestaan een herhaalrecept apart te declareren. De vergoeding hiervoor is opgenomen in andere prestatie. Het kan echter zo zijn dat een huisarts bij het verstrekken van een herhaalrecept toch een prestatie in rekening brengt, bijvoorbeeld een telefonisch consult, om zo hetzelfde bedrag vergoed te krijgen wat hij eerder ook al vergoed kreeg.
HZ09	Het in rekening brengen van meer prestaties dan geleverd	Upcoding	%meerdere consulten per patiënt op dezelfde dag	T2	Een huisarts kan bij één consument ten onrechte meerdere consulten in rekening brengen terwijl feitelijk sprake is van één patiëntencontact. Dit is niet toegestaan, ook niet op het moment dat er in één consult meerdere zorgvragen worden behandeld.
HZ10	Het in rekening brengen van meer prestaties dan geleverd	Upcoding	% (consult & M&I verrichting) per patiënt op dezelfde dag	T3	Een huisarts kan bij één consument ten onrechte zowel een consult als een M&I verrichting in rekening brengen terwijl feitelijk sprake is van één patiëntencontact. Dit is niet toegestaan, aangezien het consult verdisconteerd zit in het tarief van de M&I verrichting. In feite wordt op dat moment de eenmaal geleverde zorg dubbel gedeclareerd.
HZ11	Het op meerdere manieren in rekening brengen van ketenzorg	Upcoding	% (keten-DBC & M&I verrichting) per patiënt op dezelfde dag	T3	Een patiënt komt voor een regulier contact binnen de diabetes keten-DBC bij de huisarts. De huisarts declareert deze zorg vervolgens zowel als onderdeel van de keten-DBC als dat hij een M&I verrichting in rekening brengt.
HZ12	Het onterecht in rekening brengen van een ANW-toeslag	Upcoding	%ANZ toeslagen	T3	In rekening brengen van ANW-tarief aan de randen van de overgang van reguliere zorg naar ANW. Bijvoorbeeld een huisartsenpraktijk die tot 18.00 uur open is, maar de zorg vanaf 17.00 als ANW in rekening brengt.
HZ13	Het onterecht in rekening brengen van een ANW-toeslag	Spookzorg	%ANZ toeslagen (op weekdagen)	T3	Het in rekening brengen van een ANW-tarief voor een regulier avondspreekuur.
HZ14	Het onterecht in rekening brengen van een ANW-toeslag	Spookzorg	%ANZ toeslagen op niet erkende feestdagen	T3	In rekening brengen van ANW overdag op niet erkende feestdagen, zoals carnaval.

Mondzorg

Nr.	Scenario	Fraude categorie	Kenmerken	Type	Voorbeeld
MZ01	Het leveren van zorg terwijl de machtiging ontbreekt.	Onverzekerde zorg	Buiten scope: materiële controle	N.v.t.	Een patiënt komt bij de tandarts voor implantologie. Er is hiervoor geen indicatie gesteld. Na behandeling wordt echter ingevuld dat deze indicatie er wel was zodat de zorg alsnog wordt vergoed.
MZ02	Het leveren van meer uitgebreide/duurdere zorg dan nodig.	Overbehandelen	%kronen (meerdere categorieën)	T4	Een patiënt komt bij de tandarts vanwege tandbederf. Gezien de ernst hiervan kan de tandarts volstaan met een tweevlaksvulling. De tandarts kiest er echter voor een kroon te plaatsen. De tandarts lost hiermee het probleem op, maar levert duurdere zorg dan nodig is.
MZ03	Het vaker leveren van zorg dan nodig.	Overbehandelen	gemiddelde periode tussen röntgenfoto's (of andere behandelingen)	T2	"Voor een goede controle van het gebit maken tandartsen ongeveer elke 2 jaar een röntgenfoto.
MZ04	Het in rekening brengen van zorg die niet geleverd is	Spookzorg	#behandelingen per patiënt #behandeldatums per patiënt	T2	Bij een patiënt worden echter elk jaar röntgenfoto's genomen, zonder dat de gezondheid van het gebit daar aanleiding toe geeft."
MZ05	Het in rekening brengen van zorg die deels niet geleverd is	Spookzorg	# behandelingen per patiënt € per behandeling	T2	Een patiënt komt bij de tandarts voor het trekken van een tand. Naast het in rekening brengen van het trekken van deze tand stuurt de tandarts nog 3 facturen, elk op een andere datum, voor het trekken van tanden bij hetzelfde kind. Deze zorg is echter nooit geleverd.
MZ06	Het declareren van duurdere behandelingen	Upcoding	%gebitsreiniging uitgebreid	T3	Een patiënt komt bij de tandarts. De tandarts voert enkel een controle uit. Hij brengt vervolgens zowel een controle als het verwijderen van tandsteen in rekening bij de zorgverzekeraar.
MZ07	Het declareren van duurdere behandelingen	Upcoding	#drievlaksvulling (meerdere categorieën)	T4	Een patiënt komt bij de tandarts voor een vulling. Hij krijgt een éénvlaksvulling. De tandarts brengt echter een drievlaksvulling in rekening.
MZ08	Het meerdere malen indienen van dezelfde factuur	Dubbel claimen	Buiten scope: materiële controle	N.v.t.	Bij een reguliere controle verwijdert een tandarts tandsteen. Ongeacht de hoeveelheid tandsteen brengt deze tandarts altijd de prestatie gebitsreiniging uitgebreid in rekening. Hij zou echter bij het beperkt aanwezig zijn van tandsteen de prestatie gebitsreiniging beperkt in rekening moeten brengen.
MZ09	Het meerdere malen indienen van dezelfde factuur	Dubbel claimen	Buiten scope: materiële controle	N.v.t.	Een tandarts brengt de geleverde zorg in rekening bij zijn patiënt. Deze betaalt en stuurt de factuur naar zijn verzekeraar. De aanbieder stuurt de factuur echter ook naar de verzekeraar. In principe zal op basis van de controle van de verzekeraar slechts 1 factuur worden betaald. En het is de bedoeling dat dit de door de patiënt ingediende factuur is. De tandarts handelt niet juist.
MZ10	Op naam van een andere patiënt declareren.		Buiten scope: patiënt	N.v.t.	Een tandarts factureert via een factoringmaatschappij. Hij heeft echter al maanden geen geld ontvangen vanwege de slechte financiële positie van de factoringmaatschappij. Daarom besluit de tandarts zijn facturen van de afgelopen drie maanden ook nog rechtstreeks bij de verzekeraar in rekening te brengen. Hier bestaat het risico dat dezelfde factuur twee keer wordt betaald.
MZ11	Op naam van een andere patiënt declareren.		Buiten scope: patiënt	N.v.t.	Een patiënt weet dat hij voor de behandeling die hij zal ondergaan zelf moet gaan betalen. Zijn neef is echter wel verzekerd en hij is maar 2 jaar jonger en lijkt veel op de patiënt. De patiënt neemt de zorgpas van zijn neef mee. De tandarts heeft dit niet in de gaten en brengt zo ongemerkt de zorg in rekening bij de verzekeraar op naam van de neef waardoor er onterecht wordt vergoed
MZ12	Het splitsen van de factuur	Opknippen	Buiten scope: patiënt & zorgaanbieder	N.v.t.	Een patiënt moet een dure behandeling ondergaan die hij zelf moet betalen. Hij spreekt echter met zijn tandarts af dat de behandeling niet op zijn naam, maar op naam van zijn zoon van 12 wordt gedaan. Op deze wijze betaalt de patiënt zelf

					niets (de zorg valt immers onder de basisverzekering) en is de tandarts er zeker van dat de factuur voor de geleverde zorg wordt betaald (aangezien de verzekeraar naar verwachting eerder betaalt dan de patiënt).

Medisch specialistische zorg

Nr.	Scenario	Fraude categorie	Kenmerken	Type	Voorbeeld
MS01	Een andere zorgactiviteit vastleggen, die wel verzekerd is	Onverzekerde zorg	%onverzekerde zorg	T2 (T1)	Bij een patiënt wordt een circumcisie uitgevoerd. Dit is onverzekerde zorg. Er wordt een andere zorgactiviteit (phrenulumplastiek) geregistreerd, omdat dit wel verzekerde zorg is.
MS02	De medisch indicatie wordt niet volgens de Zvw ingevuld	Onverzekerde zorg	%onverzekerde zorg	T2 (T1)	Patiënt ondergaat een spataderoperatie. De medisch specialist geeft aan dat het gaat om verzekerde aanspraak, terwijl het op basis van de Zvw een cosmetische operatie betreft. Er wordt dan een verzekerd in plaats van een onverzekerd zorgproduct afgeleid.
MS03	Declareren niet aangevraagd onderzoek.	Overbehandelen	Buiten scope: enquête / evaluatie	N.v.t.	De huisarts doet een aanvraag aan een eerstelijns diagnostisch centrum voor een bloedonderzoek. Het diagnostische centrum voert vervolgens meer onderzoeken uit dan aangevraagd. Ook komt het voor dat de consument meer onderzoeken aankruist op het formulier. In beide gevallen declareert het diagnostisch centrum meer dan op basis van de oorspronkelijke aanvraag van de huisarts.
MS04	Verkeerd vastleggen van zorgactiviteiten.	Spookzorg	%Poliklinische wondexcisie en wondtoilet zonder verwijzing vastgelegd (meerdere)	T2 (T1)	Patiënt komt op SEH met kleine hoofdwond. De chirurg kijkt naar de wond en plakt een pleister. Vervolgens wordt de zorgactiviteit 'Poliklinische wondexcisie en wondtoilet zonder verwijzing vastgelegd'. Onder wondexcisie en wondtoilet wordt verstaan locaalanesthesie, inspectie, reiniging, excisie en/of hechting van de wond(en). De beschrijving van deze zorgactiviteit sluit niet aan bij de daadwerkelijk geleverde zorg.
MS05	Eén uitgevoerde zorgactiviteit vaker registreren.	Spookzorg	%zwaarder product	T2 (T1)	Interventiecardioloog en cardiochirurg voeren gezamenlijk AICD-implantatie (=inwendige defibrillator bij hartritmestoornissen) uit. Beiden registreren een operatieve zorgactiviteit, terwijl de operatie maar één keer is uitgevoerd. Het gevolg is dat een afleiding onterecht plaatsvindt naar een zwaarder product.
MS06	Een verkeerde koppeling van lokale verrichtingen (CBV-codering) naar zorgactiviteiten.	Spookzorg	%zorgactiviteitcode	T2 (T1)	"Ziekenhuis koppelt interne code 'verwijderen oorsmeer' aan de zorgactiviteitcode 031712. (verwijdering uit de gehoorgang van een of meerdere poliepen of corpora aliena). Dit leidt af naar een zwaarder product. Er is geen aparte zorgactiviteitcode voor 'verwijderen oorsmeer', maar dit valt binnen een regulier consult. Ziekenhuizen hanteren vaak het uitgangspunt dat elke interne code aan een zorgactiviteitcode gekoppeld moet worden.
MS07	Onterechte toepassing verlengde arm.	Spookzorg	Buiten scope: materiële controle	N.v.t.	De wondverpleegkundige houdt een zelfstandig spreekuur en registreert voor alle patiënten een polikliniekbezoek en een DBC. Het is niet toegestaan dat een verpleegkundige een DBC-zorgproduct opent, of dit namens de specialist doet, terwijl de specialist zelf geen patiëntencontact heeft.
MS08	Dubbele vergoeding voor kosten op het grensvlak van de eerste en tweedelijn.	Dubbele bekostiging	Buiten scope: materiële controle	N.v.t.	Een voorbeeld is de situatie waarin de huisarts bloed afneemt. De huisarts kan hiervoor een consult in rekening brengen, waarna het bloed naar een diagnostisch centrum wordt gestuurd voor verdere analyse. Het diagnostisch centrum brengt vervolgens een ordertarief in rekening, waar ook een vergoeding is opgenomen voor bloedafname terwijl het diagnostisch centrum hier geen kosten voor maakt.

MS09	Een nieuw zorgtraject openen bij een bestaande zorgvraag van een patiënt door één specialist.	Opknippen	% nieuw zorgproducten per patiënt per jaar	T2 (T1)	Zwangerschapsbegeleiding bij patiënt met hypertensie. Controle/behandeling is beëindigd en patiënt wordt terugverwezen naar de eerstelijn. Patiënt komt tijdens dezelfde zwangerschap nogmaals bij de gynaecoloog voor zwangerschapsdiabetes. Beide zorgvragen betreffende dezelfde zwangerschap en vallen binnen hetzelfde zorgtraject. Hier is dan sprake van een repeterende zorgvraag waar geen nieuw zorgproduct voor geopend mag worden.
MS10	Een nieuw zorgtraject openen bij een bestaande zorgvraag van een patiënt door een andere specialist van een ander specialisme.	Dubbel claimen	Trigger: dermatoloog tijdens behandeling reumatoloog %patiënten met dermatoloog tijdens behandeling reumatoloog	T2 (T1)	Patiënt is poliklinisch onder behandeling bij de reumatoloog voor psoriasis. Tijdens dit traject vindt ook een consult bij de dermatoloog plaats in het kader van deze zorgvraag. De reumatoloog is hoofdbehandelaar. Het is niet toegestaan dat de reumatoloog en dermatoloog beiden een zorgtraject openen
MS11	Bij een nieuw zorgtraject een activiteit uit een bestaande zorgvraag twee keer registreren.	Dubbel claimen	#consults per patiënt (per categorie) #consults per patiënt op dezelfde dag (per categorie)	T4	Het kan voorkomen dat een medisch specialist bij één patiënt meerdere zorgvragen constateert en hiervoor twee DBC-zorgproducten opent. Uit patiëntvriendelijkheid besluiten sommige aanbieders de patiënt niet een keer terug te laten komen, maar beide zorgvragen in eenzelfde consult af te handelen. In dat geval mag er echter slechts één consult geregistreerd worden. Eén van de behandeltrajecten bevat dan geen consult. De NZa ontvangt echter signalen dat aan beide DBC's een afzonderlijk consultcontact toegekend wordt.
MS12	Het hanteren van de typeringslijst van een ander specialisme.	U-bochtconstructie	Buiten scope: materiële controle	N.v.t.	Een kind komt bij de longarts op het spreekuur met een astma. Deze diagnose staat op typeringslijst van longgeneeskunde en van kindergeneeskunde. De longarts legt op basis van de kindergeneeskundelijst vast. Dit is geen fraude, maar ongewenst gedrag. Men kan dus op een andere AGB-code, bijvoorbeeld een reumatoloog die de AGBcode kindergeneeskunde hanteert, de duurdere variant zorg declareren. Dat is niet in de geest met het beoogde doel van de typeringslijsten.
MS13	Inplannen van verrichtingen na het afsluitmoment.	Opknippen	# klinische subtrajecten per patiënt (alleen patiënten met minimaal 1 subtraject)	T2	De subtrajecten kennen vaste afsluitmomenten. Bij een klinisch subtraject is het afsluitmoment de 42e dag na ontslagdatum. Dit geeft de behandelaar de prikkel om deze grens te overschrijden, zodat nieuwe dagverpleging of een nacontrole in een nieuw subtraject terecht komt. In feite kan de specialist dan een nieuw subtraject openen.
MS14	Afsluitgrenzen bij conservatieve behandelingen.	Opknippen	# niet-klinische subtrajecten per patiënt (alleen patiënten met minimaal 1 niet-klinische subtraject)	T2	De subtrajecten kennen vaste afsluitmomenten. Bij een niet-klinisch subtraject met conservatieve behandelingen gelden aparte afsluitregels. In dat geval mag bijvoorbeeld bij een initieel subtraject op de 90e dag na opening het subtraject gesloten worden. Dit geeft de behandelaar de prikkel om deze grens te overschrijden, zodat een nieuwe conservatieve behandeling in een nieuw subtraject terecht komt. In feite kan de specialist dan een nieuw zorgproduct openen.
MS15	Instellingen die een klinisch product declareren in plaats van een poliklinisch product, doordat een verpleegdag wordt geregistreerd (al dan niet daadwerkelijk geleverd volgens de definitie van een verpleegdag).	Spookzorg	# klinische korte opnames /patiënt	T2	Een voorbeeld is de klinische korte opnames. Dit is een aandachtsgebied voor verzekeraars in controles, omdat een enkele verpleegdag direct leidt tot een klinisch product. Ziekenhuizen mogen pas een verpleegdag in rekening brengen als de patiënt er overnacht, waarbij de patiënt is opgenomen voor 12 uur en pas is na 7 uur 's ochtends wordt ontslagen. Dit betekent dat bij een opname na twaalven de eerste 24 uur geen verpleegdag geregistreerd mag worden, terwijl daar mogelijk wel kosten tegenover staan.

Fysiotherapie

Nr.	Scenario	Fraude categorie	Kenmerken	Type	Voorbeeld
FY01	Het niet voldoen aan de eisen van een lange zitting	Upcoding	%lange zitting	T2 (T1)	De lange zitting is bedoeld voor patiënten met complexe en/of meervoudige zorgvragen. De aandoening en de situatie van de patiënt brengen met zich mee dat het niet mogelijk is om de interventie in een reguliere zitting uit te voeren. Deze prestatie kan door de zorgaanbieder alleen in rekening worden gebracht indien er sprake is van een complex behandelprogramma.
FY02	Het onterecht in rekening brengen van toeslagen voor uitbehandeling	Upcoding	Buiten scope: materiële controle	N.v.t.	In het geval de zorgaanbieder de patiënt bezoekt, kan naast de zitting een toeslag voor uitbehandeling worden gedeclareerd. Dit mag niet wanneer de behandeling in een inrichting plaats vindt. Hiervoor geldt een aparte toeslag.
FY03	Dezelfde prestatie dubbel betalen	Dubbel claimen	Buiten scope: materiële controle	N.v.t.	Het indienen van een declaratie zowel door de zorgaanbieder als de patiënt. Omdat de meeste zorg via de aanvullende verzekering wordt gedekt, wordt er ook vaak een factuur naar de patiënt gestuurd. Als zowel de patiënt als de zorgaanbieder de rekening naar de zorgverzekeraars stuurt, bestaat de kans dat dezelfde prestatie dubbel wordt betaald.
FY04	Niet volgens de regels declareren	Dubbel claimen	% meerdere prestaties/behandelingen voor dezelfde patiënt op dezelfde dag	T2 (T1)	Het onterecht twee of meer dezelfde prestaties/behandelingen of een combinatie van twee of meer verschillende prestaties/behandelingen verricht op 1 dag declareren. Ook al kosten sommige patiënten dan misschien meer tijd, zorgaanbieders mogen niet zomaar meerdere consulten declareren.
FY05	Niet volgens de regels declareren	Dubbel claimen	%Inrichtingstoeslagen per patiënt per jaar	T2 (T1)	Het ten onrechte in rekening brengen van (inrichtings-)toeslagen. Inrichtingstoeslagen mogen alleen in rekening worden gebracht als er sprake is van een incidentele behandeling waarvoor de zorgaanbieder de praktijk moet verlaten. Als de locatie waar behandeld wordt een meer permanent karakter heeft (zorgaanbieder is een vast (dag)deel per week op een vaste behandelplek in de inrichting), dan is de inrichtingstoeslag niet van toepassing. In de praktijk blijkt er nog wel eens onduidelijkheid te bestaan over de begrippen ‘permanent karakter’, ‘vast (dag)deel’ en ’vaste behandelplek’ en wordt hier verkeerd gedeclareerd.
FY06	Declareren van manuale therapie, terwijl fysiotherapie is geleverd	Upcoding	%manuale therapie	T2 (T1)	Het betreft een gewone onafgebroken zitting, waarin de patiënt wordt behandeld, begeleid en/of geadviseerd voor één of meer indicaties.
FY07	Kinderfysiotherapie in plaats van gewone therapie declareren	Upcoding	%Kinderfysiotherapie	T2 (T1)	Het betreft een gewone onafgebroken zitting gericht op kinderen, waarin de patiënt wordt behandeld, begeleid en/of geadviseerd voor één of meer indicaties.
FY08	Het in rekening brengen van individuele behandelingen terwijl er sprake is van groepsbehandeling	Upcoding	#dagen met gelijke behandelingen (type/duur) >= 12 uur %groeps- versus individuele behandeling Enquête / Evaluatie	T2 (T1)	Groepszittingen kunnen in verschillende vormen naar voren komen. In alle gevallen moet er aan een aantal voorwaarden worden voldaan. Zo gelden er eisen voor de duur van de behandeling, de indicatiestelling ‘groepsbehandeling’ gebeurt in overleg met de patiënt en de individuele behandelplannen worden uitgebreid met een groepsbehandelplan.
FY09	Het declareren van een telefonische zitting die niet heeft plaatsgevonden	Spookzorg	Buiten scope: enquête / evaluatie	N.v.t.	Voor het in rekening brengen van een telefonische zitting moet aan verschillende voorwaarden zijn voldaan. Zo dient deze onder andere ter vervanging van een reguliere zitting te zijn en plaats te vinden tijdens of kort na een behandelepisode. Tevens dient de patiënt uitdrukkelijk te zijn geïnformeerd over het doel van de telefonische zitting.
FY10	Het declareren van een huisbezoek, wat niet heeft plaatsgevonden	Spookzorg	%huisbezoek verhoudingen tussen telefonische zittingen en huisbezoeken enerzijds en gewone	T3	Ook voor het in rekening brengen van een huisbezoek (uitbehandeling) gelden bepaalde voorwaarden. Het gaat hier om of een huisbezoek daadwerkelijk heeft plaatsgevonden of niet.

			zittingen en praktijk behandelingen anderzijds. Enquête / Evaluatie		
FY11	Het declareren van zorg terwijl er niet wordt gewerkt	Spookzorg	Buiten scope: enquête / evaluatie	N.v.t.	Een aanbieder brengt voor 5 dagen per week zorg in rekening terwijl bekend is dat hij maar 4 dagen werkt.
FY12	Het declareren van andere prestaties dan wat er is geleverd	Spookzorg	Buiten scope: enquête / evaluatie	N.v.t.	Een praktijk fysiotherapie verspreidt flyers met daarop reclame voor fietsmetingen / afstellen waarvan de kosten € 190 zijn. Daarbij wordt vermeld dat op het moment dat degene die deze fietsmeting afneemt ook lichamelijk klachten heeft (en voldoende is verzekerd) de fysiotherapeut dit gedeeltelijk op de zorgverzekering kan verhalen. De kosten voor de patiënt bedragen dan € 50 en verder worden er vijf behandelingen gedeclareerd bij de zorgverzekeraar om zo de resterende € 140 te voldoen. Er worden dus behandelingen gedeclareerd die niet plaatsvinden. In genoemd voorbeeld zullen deze ten laste van de aanvullende verzekering worden gebracht.

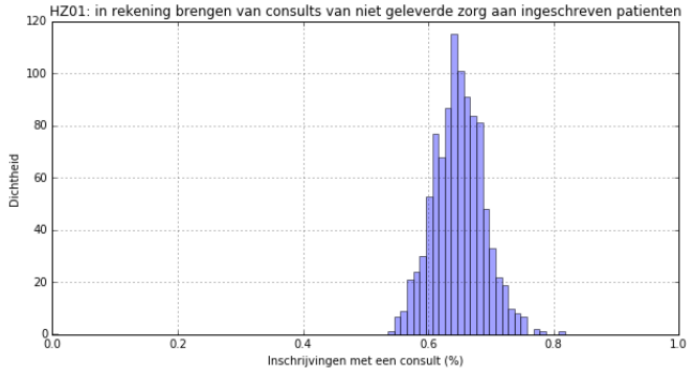
Bijlage 2: Jupyter code voor generatie test datasets

T1: Normale verdeling met één kenmerk

In [110]:
```
n_samples = 1000
n_outliers = n_samples // (100//5) # 5%
n_noise =n_samples // (100//2)   # 2%
```

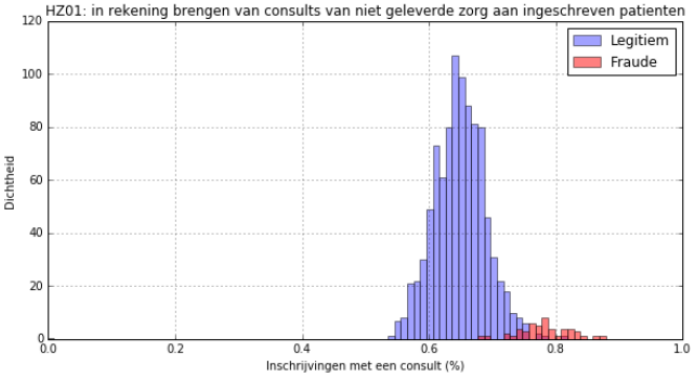
HZ01: in rekening brengen van consults van niet geleverde zorg aan ingeschreven patienten

In [130]:
```
bins = np.linspace(0,1,100)
np.random.seed(2)
x = np.random.normal(0.65, 0.04, n_samples)
plt.hist(x, bins=bins, normed=False, color='#4444ff', alpha=0.5)
plt.grid(True)
plt.title('HZ01: in rekening brengen van consults van niet geleverde zorg aan ingeschreven patienten')
plt.xlabel('Inschrijvingen met een consult (%)')
plt.ylabel('Dichtheid')
plt.xlim(0,1)
plt.savefig("T1_norm.png")
plt.show()
```



Outliers toevoegen ¶

In [131]:
```
np.random.seed(0)
# Add outliers
true_outliers = np.sort(np.random.randint(0, len(x), n_outliers))
inliers = list(set(range(len(x))).difference(true_outliers))
x[true_outliers] = np.random.normal(0.78, 0.04, n_outliers)
#x[true_outliers] = np.random.random_sample(n_outliers) * .2 + 0.8
plt.hist(x[inliers], bins=bins, normed=False, color='#4444ff', alpha=0.5, label='Legitiem')
plt.hist(x[true_outliers], bins=bins, normed=False, color="red", alpha=0.5, label='Fraude')
plt.grid(True)
plt.title('HZ01: in rekening brengen van consults van niet geleverde zorg aan ingeschreven patienten')
plt.xlabel('Inschrijvingen met een consult (%)')
plt.ylabel('Dichtheid')
plt.xlim(0,1)
plt.legend(loc='upper right')
plt.savefig("T1_outl.png")
plt.show()
```

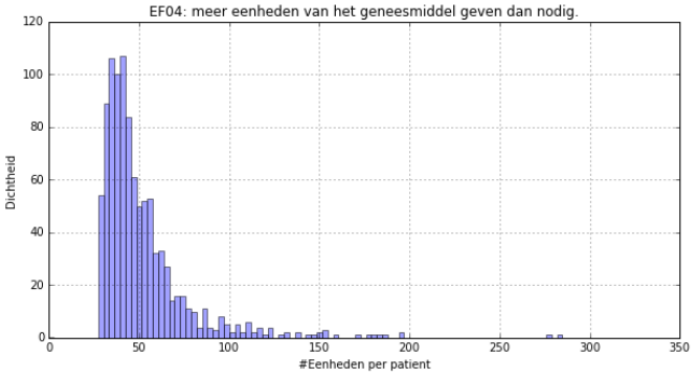


T2: Lognormale verdeling met één kenmerk

In [89]:
```
n_samples = 1000
n_outliers = n_samples // (100//5) # 5%
n_noise =n_samples // (100//2)   # 2%
```

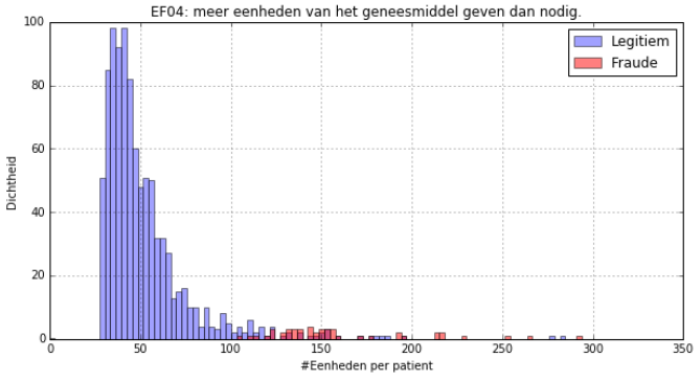
HZ01: in rekening brengen van consults van niet geleverde zorg aan ingeschreven patienten

In [148]:
```
bins = np.linspace(0,300,100)
np.random.seed(2)
mu, sigma = 3., 0.8 # mean and standard deviation
x = np.random.lognormal(mu, sigma, n_samples)
x = x + 25
plt.hist(x, bins=bins, normed=False, color='#4444ff', alpha=0.5)
plt.grid(True)
plt.title('EF04: meer eenheden van het geneesmiddel geven dan nodig.')
plt.xlabel('#Eenheden per patient')
plt.ylabel('Dichtheid')
#plt.xlim(0,1)
plt.savefig("T2_norm.png")
plt.show()
```



Outliers toevoegen

In [149]:
```
np.random.seed(0)
# Add outliers
true_outliers = np.sort(np.random.randint(0, len(x), n_outliers))
true_inliers = list(set(range(len(x))).difference(true_outliers))

x[true_outliers] = x[true_outliers] * (2 + 3 * np.random.random_sample())
#x[true_outliers] = np.random.random_sample(n_outliers) * .2 + 0.8
plt.hist(x[true_inliers], bins=bins, normed=False, color='#4444ff', alpha=0.5, label='Legitiem')
plt.hist(x[true_outliers], bins=bins, normed=False, color="red", alpha=0.5, label='Fraude')
plt.grid(True)
plt.title('EF04: meer eenheden van het geneesmiddel geven dan nodig.')
plt.xlabel('#Eenheden per patient')
plt.ylabel('Dichtheid')
plt.xlim(0)
plt.legend(loc='upper right')
plt.savefig("T2_outl.png")
plt.show()
```

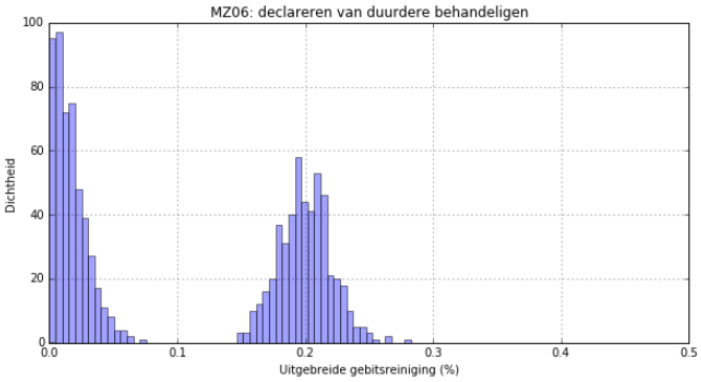


T3: Gecombineerde normale verdeling met één kenmerk

In [133]:
```
n_samples = 1000
n_outliers = n_samples // (100//5) # 5%
n_noise =n_samples // (100//2)   # 2%
```

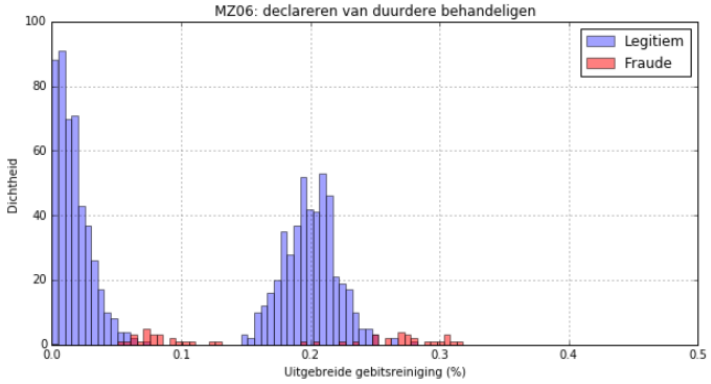
%Uitgebreide gebitsreiniging

In [136]:
```
bins = np.linspace(0,0.5,100)
np.random.seed(2)
x = np.concatenate([np.random.normal(0.2,0.02, 500),
                    np.random.normal(0.01, 0.02, 500)])
x = abs(x);
plt.hist(x, bins=bins, normed=False, color='#4444ff', alpha=0.5)
plt.xlim(0,0.5);
plt.grid(True)
plt.title('MZ06: declareren van duurdere behandeligen')
plt.xlabel('Uitgebreide gebitsreiniging (%)')
plt.ylabel('Dichtheid')
plt.savefig("T3_norm.png")
plt.show()
```



Outliers toevoegen

In [194]:
```
np.random.seed(0)
#np.random.uniform(low=-6, high=6, size=(n_outliers, 2)


# Add outliers
true_outliers = np.sort(np.random.randint(0, len(x), n_outliers))
inliers = list(set(range(len(x))).difference(true_outliers))
x[true_outliers] =  np.concatenate([np.random.normal(0.08, 0.02, n_outliers // 2),
                                    np.random.normal(0.27, 0.03, n_outliers // 2)])
plt.hist(x[inliers], bins=bins, normed=False, color='#4444ff', alpha=0.5, label='Legitiem')
plt.hist(x[true_outliers], bins=bins, normed=False, color="red", alpha=0.5, label='Fraude')
plt.xlim(0,0.5);
plt.grid(True)
plt.title('MZ06: declareren van duurdere behandeligen')
plt.xlabel('Uitgebreide gebitsreiniging (%)')
plt.ylabel('Dichtheid')
plt.legend(loc='upper right')
plt.savefig("T3_out.png")
plt.show()
```

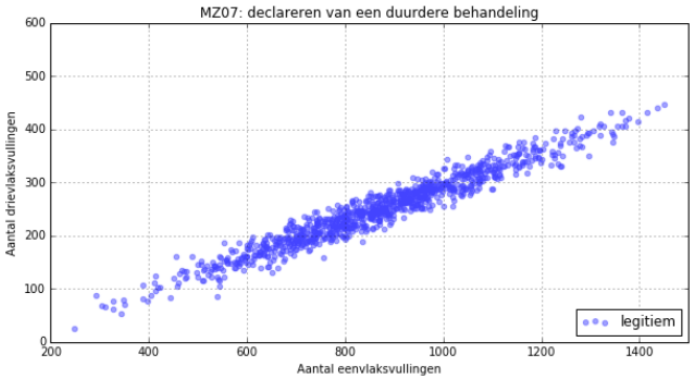


T4: Meerdere afhankelijke kenmerken (of categorieën)

In [3]:
```
n_samples = 1000
n_outliers = n_samples // (100//5) # 5%
n_noise =n_samples // (100//2)   # 2%
```

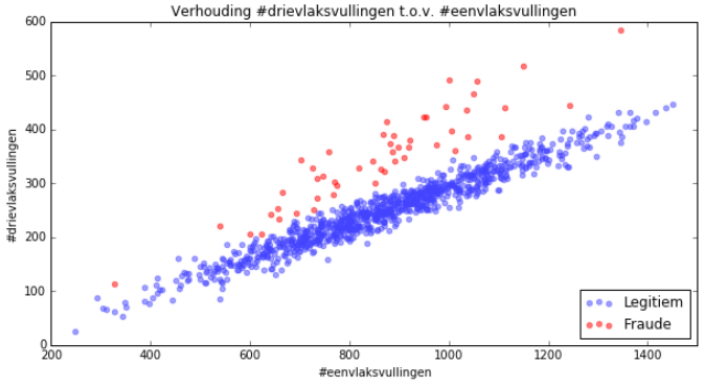
MZ07: declareren van een duurdere behandeling

In [4]:
```
# eenvlaksvulling  : 50..1500
# drievlaksvulling : 5..500

np.random.seed(5)
xo, yo, coef = datasets.make_regression(n_samples=n_samples, n_features=1,n_informative=1, noise=20,
                                        coef=True, random_state=0)
# eenvlaksvulling
xo = (xo + 4)/7 * 1450 + 50
# drievlaksvulling
yo = (yo + 300)/600  * 495 + 5
plt.scatter(xo, yo, color='#4444ff', marker='o', alpha=0.5, label='legitiem')
plt.legend(loc='lower right')
plt.xlim(200, 1500)
plt.ylim(0, 600)
plt.grid(True)
plt.title('MZ07: declareren van een duurdere behandeling')
plt.xlabel('Aantal eenvlaksvullingen')
plt.ylabel('Aantal drievlaksvullingen')
plt.savefig("T4_norm.png")
plt.show()
```



Outliers toevoegen

In [240]:
```
# Add outlier data
x = copy.copy(xo)
y = copy.copy(yo)
true_outliers = np.sort(np.random.randint(0, len(x), n_outliers))
inliers = list(set(range(len(x))).difference(true_outliers))
y[true_outliers] = y[true_outliers] * (1.2+np.random.random(n_outliers) * 0.4)
plt.scatter(x[inliers], y[inliers], color='#4444ff', marker='o', alpha=0.5, label='Legitiem')
plt.scatter(x[true_outliers], y[true_outliers], color='red', marker='o', alpha=0.5, label='Fraude')
plt.legend(loc='lower right')
plt.xlim(200, 1500)
plt.ylim(0, 600)
plt.xlabel('#eenvlaksvullingen')
plt.ylabel('#drievlaksvullingen')
plt.title('Verhouding #drievlaksvullingen t.o.v. #eenvlaksvullingen')
plt.savefig("T4_out.png")
plt.show()
```



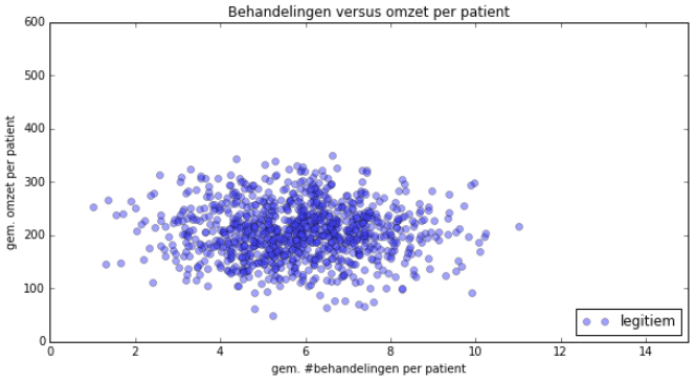
T5: Eén cluster met twee onafhankelijke kenmerken

In [631]:
```
n_samples = 1000
n_outliers = n_samples // (100//5) # 5%
n_noise =n_samples // (100//2)   # 2%
```

GG07: Behandelingen versus omzet per patient

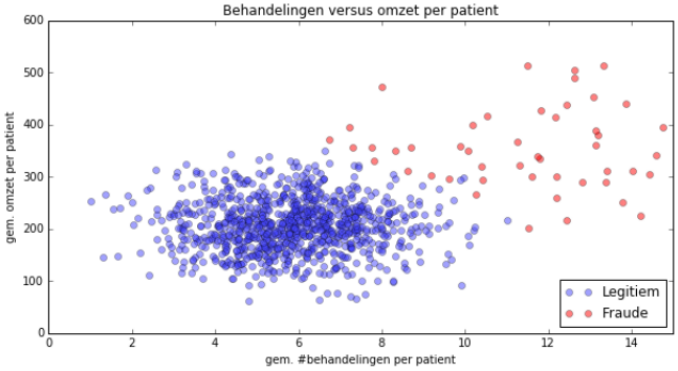
In [632]:
```
centers = [[1, 1]]
x, labels_true = make_blobs(n_samples=n_samples, centers=centers, cluster_std=0.1, random_state=0)
x = StandardScaler().fit_transform(x)

x[:, 0] = x[:, 0] - min(x[:, 0])
x[:, 0] = x[:, 0] / max(x[:, 0])
x[:, 0] = x[:, 0] * 10 + 1

x[:, 1] = x[:, 1] - min(x[:, 1])
x[:, 1] = x[:, 1] / max(x[:, 1])
x[:, 1] = x[:, 1] * 300 + 50
```

In [633]:
```
plt.plot(x[:, 0], x[:, 1], 'o', markerfacecolor='#4444ff', markeredgecolor='k', markersize=6, alpha=0.5, label='legitier
plt.xlim(0,15);
plt.ylim(0,600);
plt.xlabel('gem. #behandelingen per patient')
plt.ylabel('gem. omzet per patient')
plt.title('Behandelingen versus omzet per patient')
plt.legend(loc='lower right')
plt.savefig("T5_norm.png")
plt.show()
```



Outliers toevoegen

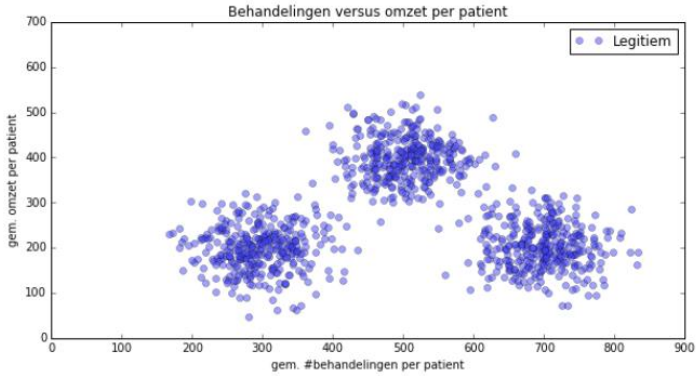
In [634]:
```
true_outliers = np.sort(np.random.randint(0, len(x), n_outliers))
x[true_outliers, 0] = x[true_outliers, 0] + 3 + 5 * deltaX
x[true_outliers, 1] = x[true_outliers, 1] + 50 + 200 * deltaY
inliers = list(set(range(len(x))).difference(true_outliers))
plt.plot(x[inliers, 0], x[inliers, 1], 'o', markerfacecolor='#4444ff', markeredgecolor='k', markersize=6, alpha=0.5, lak
plt.plot(x[true_outliers, 0], x[true_outliers, 1], 'o', markerfacecolor='red', markeredgecolor='k', markersize=6, alpha=
plt.xlim(0,15);
plt.ylim(0,600);
plt.xlabel('gem. #behandelingen per patient')
plt.ylabel('gem. omzet per patient')
plt.title('Behandelingen versus omzet per patient')
plt.legend(loc='lower right')
plt.savefig("T5_out.png")
plt.show()
```



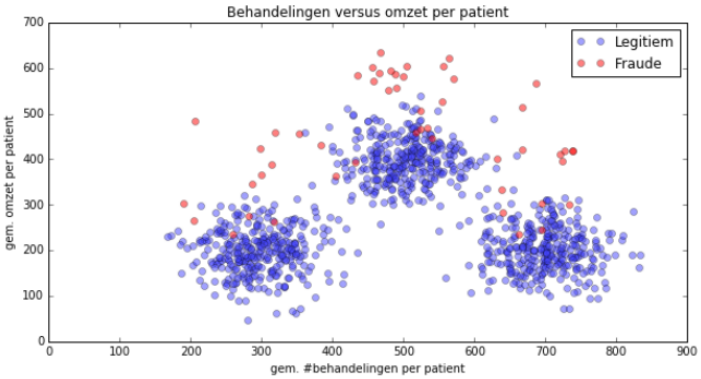
T6: Meerdere clusters met twee onafhankelijke kenmerken

In [2]:
```
n_samples = 1000
n_outliers = n_samples // (100//5) # 5%
n_noise =n_samples // (100//2)   # 2%
```

In [3]:
```
centers = [[300, 200], [500, 400], [700, 200]]
x, labels_true = make_blobs(n_samples=n_samples, centers=centers, cluster_std=50, random_state=0)
#x = StandardScaler().fit_transform(x)

plt.plot(x[:, 0], x[:, 1], 'o', markerfacecolor='#4444ff', markeredgecolor='k', markersize=6, alpha=0.5, label='Legitiem
plt.xlim(0,900);
plt.ylim(0,700);
plt.xlabel('gem. #behandelingen per patient')
plt.ylabel('gem. omzet per patient')
plt.title('Behandelingen versus omzet per patient')
plt.legend(loc='upper right')
plt.savefig("T5_norm.png")
plt.show()
```



In [4]:
```
true_outliers = np.sort(np.random.randint(0, len(x), n_outliers))
x[true_outliers, 0] = x[true_outliers, 0] + 3 + 5 * np.random.random(size=(n_outliers))
x[true_outliers, 1] = x[true_outliers, 1] + 50 + 200 * np.random.random(size=(n_outliers))
inliers = list(set(range(len(x))).difference(true_outliers))
plt.plot(x[inliers, 0], x[inliers, 1], 'o', markerfacecolor='#4444ff', markeredgecolor='k', markersize=6, alpha=0.5, lab
plt.plot(x[true_outliers, 0], x[true_outliers, 1], 'o', markerfacecolor='red', markeredgecolor='k', markersize=6, alpha=
plt.xlim(0,900);
plt.ylim(0,700);
plt.xlabel('gem. #behandelingen per patient')
plt.ylabel('gem. omzet per patient')
plt.title('Behandelingen versus omzet per patient')
plt.legend(loc='upper right')
plt.savefig("T5_out.png")
plt.show()
```

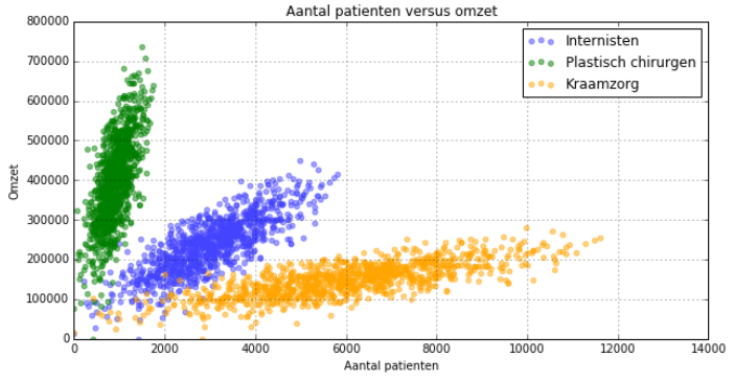


T7: Meerdere categorieën van afhankelijke variabelen

In [2]:
```
n_samples = 1000
n_noise =n_samples // (100//2)   # 2%
```

MZ07: declareren van een duurdere behandeling

In [3]:
```
np.random.seed(5)
x1, y1, coef = datasets.make_regression(n_samples=n_samples, n_features=1, n_informative=1, noise=60,coef=True, random_s
x2, y2, coef = datasets.make_regression(n_samples=n_samples, n_features=1, n_informative=1, noise=90,coef=True, random_s
x3, y3, coef = datasets.make_regression(n_samples=n_samples, n_features=1, n_informative=1, noise=70,coef=True, random_s
x1 = (x1 - min(x1)) * 1000
y1 = (y1 - min(y1)) * 700
x2 = (x2 - min(x2)) * 300
y2 = (y2 - min(y2)) * 900
x3 = (x3 - min(x3)) * 2000
y3 = (y3 - min(y3)) * 400
plt.scatter(x1, y1, color='#4444ff', marker='o', alpha=0.5, label='Internisten')
plt.scatter(x2, y2, color='green', marker='o', alpha=0.5, label='Plastisch chirurgen')
plt.scatter(x3, y3, color='orange', marker='o', alpha=0.5, label='Kraamzorg')
plt.legend(loc='upper right')
plt.xlim(0)
plt.ylim(0)
plt.grid(True)
plt.title('Aantal patienten versus omzet')
plt.xlabel('Aantal patienten')
plt.ylabel('Omzet')
plt.savefig("T7.png")
plt.show()
```



Bijlage 3: CSV export van test datasets

T1: Normale verdeling met één kenmerk

t1.csv

T2: Lognormale verdeling met één kenmerk

t2.csv

T3: Gecombineerde normale verdeling met één kenmerk

t3.csv

T4: Meerdere afhankelijke kenmerken (of categorieën)

t4.csv

T5: Eén cluster met twee onafhankelijke kenmerken

t5.csv

T6: Meerdere clusters met twee onafhankelijke kenmerken

t6.csv